

Pred. No. is the number of results predicted by chance to have a

Copyright (c) 1993 - 2004 Compugen Ltd.
 GenCore version 3.1.6

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	1640.6	99.5	2076	8	AY006482	Artemisia
2	1637.4	99.3	2076	6	BD227440	Transgeni
3	1635.8	99.2	2080	8	AF138959	Artemisia
4	1634.2	99.1	2106	8	AA2251751	Artemisia
5	1626.2	98.6	2067	8	AF327526	Artemisia
6	857	52.0	1902	8	AF472361	Artemisia
7	836	50.7	2026	8	AA271793	Artemisia
8	776.2	47.1	1976	8	SCA304452	Solidago
9	723.4	43.9	1946	8	RA7ERPENS	Artemisia
10	723.4	43.9	1985	8	AF157059	Artemisia
11	723.2	43.9	4392	8	AF327527	Artemisia
12	721.8	43.8	1893	8	AA249561	Artemisia
13	588	35.7	1886	8	AF304444	Artemisia
14	559	33.9	1843	8	AA2271792	Artemisia
15	510.6	31.0	538	6	BD227439	Transgeni
16	424.6	25.7	1890	8	GACADIC2	Gossypium a
17	424.6	25.7	1893	8	GAU323205	Gossypium a
18	418.2	25.4	1820	8	AF270425	Gossypium a
19	418.2	25.4	1892	8	GAU232026	Gossypium a
20	416.6	25.3	1851	8	GHU98318	Gossypium h
21	399.6	24.2	1883	8	GAU27535	Gossypium a
22	395.6	24.2	1935	8	GACADSYNH	Gossypium a
23	396.2	24.0	1973	6	AX529025	Sequence
24	390.2	23.7	1900	8	AF288465	Citrus ju
25	388.4	23.6	2030	8	AF279455	Lycopersi
26	386.4	23.4	1647	8	AF441124	Citrus si
27	383.2	23.2	1877	8	AF411120	Citrus x
28	369.8	22.4	2605	6	AX529022	Sequence
29	364	22.1	1879	6	AR183920	Sequence
30	364	22.1	1879	8	AF035630	Lycopersi
31	364	22.1	2003	8	AF279454	Lycopersi
32	364	22.1	2018	8	AF279453	Lycopersi
33	363.6	22.0	1672	6	AX529019	Sequence
34	361.2	21.9	2000	8	AF279456	Lycopersi
35	357.6	21.7	2024	6	BD227674	Syntheses
36	357.6	21.7	2024	6	AR183921	Sequence
37	357.6	21.7	2024	6	AR240709	Sequence
38	357.6	21.7	2024	6	AR266984	Sequence
39	357.6	21.7	2024	6	AR316333	Sequence
40	357.6	21.7	2024	6	AR338475	Sequence
41	357.6	21.7	2024	6	AR429882	Sequence
42	357.6	21.7	2024	8	AF035631	Lycopersi
43	340	20.6	1867	8	AF374462	Citrus ju
44	330.4	20.0	1944	6	BD227667	Syntheses
45	330.4	20.0	1944	6	AR240702	Sequence

ALIGNMENTS

AY006482 2076 bp mRNA linear PLN 03-JUL-2001
Artemisia annua amorpha-4,11-diene synthase mRNA, complete cds.

Artemisia annua (sweet wormwood)

Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;

11 (bases 1 to 2076)

Wallaart, T.E., Bouwmeester, H.J., Hille, J., Poppinga, L. and

Pred. No. is the number of results predicted by chance to have a

Maïjers, N.C.
Amorpha-4,11-diene synthase: cloning and functional expression of a
key enzyme in the biosynthetic pathway of the novel antimalarial
drug artemisinin
Planta 212 (3), 460-465 (2001)
21183247
MEDLINE
FUEMED
REFERENCE
2 (bases 1 to 2076)
Wallaart, T.E. and Bouwmeester, H.J.
DNA encoding amorpha-4,11-diene synthase
Patent: European Patent Office (39202854.0) - EPO Bulletin 2000/09 ;
GenoClipp Biotechnology B.V. ;
The Netherlands ;
3 (bases 1 to 2076)
Wallaart, T.E., Bouwmeester, H.J., Hille, J., Poppinga, L. and
Maïjers, N.C.A.
Direct Submission
Submitted (09-AUG-2000) GenoClipp Biotechnology B.V., Meditech
Center, L.J. Zielstraweg 1, Groningen 9713 GX, The Netherlands
Location/Qualifiers
1. .2076
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
6. .1646
/codon_start=1
/product="amorpha-4,11-diene synthase"
/protein_id="AAF98444.1"
/db_xref="GI:9837804"
/translation="MSLITEEPPIPIANFPSPISNGDQFLIYQKVEQVQVQVNDLKK
EVRLKALDIPMKHANLLKLDIEIQRLGIPYHFERIDHALQCIETYGNWNGDR
SSWFLRMKGGYVYCDVFNKYKQKQAFKQSLANDVEGLLELYEATSMRVPGBIIL
EDALGFTSRSLISINTKDAFTNPALEIQRALQPLKRLPRIEAAQYIIPYQQQDS
HNKLLKALEFNLKSHKELSHVCKWKAEDIKNAKPLDRIVCEYFWLGSQ
YEPQYSARVPETKAVAVITLIDDTYDAYGYEELKIPTAEVBSIKLCLDPLPYMK
PIYKLFMDTYTEMEFLAKGRDLPNGCKEYFVEFVNLNMEKANEHPIPTHEK
DPVILITGGANLLTTCYLGMSDIFTKESVSWASAPPLFRYGLIGRLRLNDLMTKHA
EQKHKHSSLSLTYNNNEBYAOTLIYKEVEDWKNIDINREYLVITTKNIIPRLLMAY
IYLQCFLEVOYAGKDNFTRMDEYKHLIKSLLVPMST"

FEATURES
source
CDS
Query Match 99.5%; Score 1640.6; DB 8; Length 2076;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CATGGCACTTACAGAGAAAACCTATTTCGCCCACTTCCCACTTTCCTCCAAAGCATTTG 61
DB 5 CATGTCACTTACAGAGAAAACCTATTTCGCCCACTTCCCACTTTCCTCCAAAGCATTTG 64
QY 62 GGGAGATCAGTTTCTCATCTATCAAAACCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA 121
DB 65 GGGAGATCAGTTTCTCATCTATCAAAACCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA 124
QY 122 TGATTTAAAAAGAGTGGCGCACTACTAAAGAGCTTTGGATATTCCTATGAACA 181
DB 125 TGATTTAAAAAGAGTGGCGCACTACTAAAGAGCTTTGGATATTCCTATGAACA 184
QY 182 TGCCAAATTTGTTGAAGCTGATTGATGAATTCACGCCCTTGGATACCGTATCACTTTGA 241
DB 185 TGCCAAATTTGTTGAAGCTGATTGATGAATTCACGCCCTTGGATACCGTATCACTTTGA 244
QY 242 ACGGAGATTCATGCATTGCAATGATTTATGAACAATATGTTGATGATGATGATGATGATG 301
DB 245 ACGGAGATTCATGCATTGCAATGATTTATGAACAATATGTTGATGATGATGATGATGATG 304
QY 302 TGACCGCTCTCTTATGTTTGGTTTCCTATTCGGAAGCAAGGATATTTATGTTACATGTA 361
DB 305 TGACCGCTCTCTTATGTTTGGTTTCCTATTCGGAAGCAAGGATATTTATGTTACATGTA 364
QY 362 TGTITTCATTAACATATAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTTAATGATGT 421
DB 365 TGTITTCATTAACATATAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTTAATGATGT 424

ORIGIN

QY 422 TGAAGCTTTGCTTGTAGTTGTACGAACAACTTCTATGAGGCTACCTGGGGAGATTATATT 481
DB 425 TGAAGCTTTGCTTGTAGTTGTACGAACAACTTCTATGAGGCTACCTGGGGAGATTATATT 484
QY 482 AGAAGATGCTTCTTGGTTTACAGATCTCTCTTAGCATATGACAAAAGATGCTTTTTC 541
DB 485 AGAAGATGCTTCTTGGTTTACAGATCTCTCTTAGCATATGACAAAAGATGCTTTTTC 544
QY 542 TACAAACCCCGCTCTTTTACCGGAATACAAACGGGCACTAAAGCAACCCCTTTGAAAAG 601
DB 545 TACAAACCCCGCTCTTTTACCGGAATACAAACGGGCACTAAAGCAACCCCTTTGAAAAG 604
QY 602 GTTGCCAAAGATAGAGCGCGCGCAGTACATTCCTTTCTATCAACAAAGATTCCTATAA 661
DB 605 GTTGCCAAAGATAGAGCGCGCGCAGTACATTCCTTTCTATCAACAAAGATTCCTATAA 664
QY 662 CAAGACTTTTACTTAAACTTGTAAAGTTAGAGTTCAATTTGCTTCACTTGTGCAAGA 721
DB 665 CAAGACTTTTACTTAAACTTGTAAAGTTAGAGTTCAATTTGCTTCACTTGTGCAAGA 724
QY 722 AGAGCTCAGCCATGTGTGCAAAATGGTGAAGCTTTTCGATATCAAGAAAGCAACGACCTTG 781
DB 725 AGAGCTCAGCCATGTGTGCAAAATGGTGAAGCTTTTCGATATCAAGAAAGCAACGACCTTG 784
QY 782 TTTAAGAGATAGAAATTTGTAATGTCTTTTGGGACTAGGTTTCAAGCTATGAGCCACA 841
DB 785 TTTAAGAGATAGAAATTTGTAATGTCTTTTGGGACTAGGTTTCAAGCTATGAGCCACA 844
QY 842 GTATTCCCGGGCTAGAGTTTCTTCCAAAGCTTGTCTTATATACTCTTATAGATGA 901
DB 845 GTATTCCCGGGCTAGAGTTTCTTCCAAAGCTTGTCTTATATACTCTTATAGATGA 904
QY 902 CACTTATGATGGTATGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTGGAAG 961
DB 905 CACTTATGATGGTATGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTGGAAG 964
QY 962 GTGGTCAATTAATGCTTTAGACACATTCGAAATACATGAACCGGATATCAAAATTA 1021
DB 965 GTGGTCAATTAATGCTTTAGACACATTCGAAATACATGAACCGGATATCAAAATTA 1024
QY 1022 CATGNTATACATACAGAAATGGAAAGTTTCTTGAAGAGGAGGAGACAGATCTATT 1081
DB 1025 CATGNTATACATACAGAAATGGAAAGTTTCTTGAAGAGGAGGAGACAGATCTATT 1084
QY 1082 TAACTGCGGCAAAAGATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAAGCAAAATG 1141
DB 1085 TAACTGCGGCAAAAGATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAAGCAAAATG 1144
QY 1142 GGCAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATCTATG 1201
DB 1145 GGCAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATCTATG 1204
QY 1202 CGGTGCTAACCTGCTTACAACTTTGTTATCTTGGCATGAGTATATTCACAAAAGA 1261
DB 1205 CGGTGCTAACCTGCTTACAACTTTGTTATCTTGGCATGAGTATATTCACAAAAGA 1264
QY 1262 GTCTGTGCAATGGGCTGTCTCTGACCTCTCTTTTGTAGATCTCAGGTATCTTGTGTCG 1321
DB 1265 GTCTGTGCAATGGGCTGTCTCTGACCTCTCTCTTTTGTAGTACTCAGGTATCTTGTGTCG 1324
QY 1322 ACGCTTAAATGATCTCATGACCAAGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1381
DB 1325 ACGCTTAAATGATCTCATGACCAAGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1384
QY 1382 CTTTGAAGTTATATGAAGGATATATGTCATGAGGAGTATGTCCTCAACCTTGTGATTTA 1441
DB 1385 CTTTGAAGTTATATGAAGGATATATGTCATGAGGAGTATGTCCTCAACCTTGTGATTTA 1444
QY 1442 CAAGGAAGTAGAAGTGTGTGAAGATATAAACCAGGAGTACTCTCAACATAAAGACAT 1501
DB 1445 CAAGGAAGTAGAAGTGTGTGAAGATATAAACCAGGAGTACTCTCAACATAAAGACAT 1504
QY 1502 TCAAAGCGGCTTATTGATGGCTGTGATCTATTGTCGCCAGCTTTCTTGAAGTTCAATATGC 1561

Db 1505 TCCAAGCCGTTATTGATGGCTGTGATCTATTGTGCCAGTCTTCTTGAAGTTCATATGC 1564
QY 1562 AGGAAGGATAACTTACACGATATGGGAGAGCAATACAAAATCTCATTAAGTCTCTACT 1621
Db 1565 AGGAAGGATAACTTACACGATATGGGAGAGCAATACAAAATCTCATTAAGTCTCTACT 1624
QY 1622 CGTTTATCCTATGATATATGAGGATC 1648
Db 1625 CGTTTATCCTATGATATATGACTACC 1651

RESULT 2
BD227440 2076 bp DNA linear PAT 17-JUL-2003
LOCUS Transgenic amorph-4,11-diene synthesis.
DEFINITION BD227440
ACCESSION BD227440
VERSION BD227440.1 GI:33037210
KEYWORDS JP 2002523101-A/2.
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua

REFERENCE
AUTHORS Wallaart,T.E. and Bouwmeester,H.J.
TITLE Transgenic amorph-4,11-diene synthesis
JOURNAL Patent: JP 2002523101-A 2 30-JUL-2002;
GENCLIPP BIOTECHNOLOGY BV
COMMENT OS Artemisia annua
PN JP 2002523101-A/2
PD 30-JUL-2002
PF 27-AUG-1999 JP 2000567711
PR 27-AUG-1998 EP 98020854.0
PI THORVALD BELCO WALLAART,HENDRIK JAN BOUWMEESTER PC
C12N15/09,C12N15/09,A01H5/00,C12N1/19,C12N1/21,C12N5/10,C12N9/ PC
88,C12P5/00.
PC C12P17/18/C12N9/04,(C12N1/19,C12R1:84),(C12N1/19,C12R1:645),
PC (C12N9/04,C12R1:91),(C12N9/04,C12R1:865),(C12N9/04,C12R1:84),
PC (C12N9/04,C12R1:19),(C12N9/04,C12R1:645),(C12N9/88,C12R1:91),
PC (C12N9/88,C12R1:19),(C12N9/88,C12R1:84),(C12N9/88,C12R1:865),
PC (C12N9/88,C12R1:645),(C12N15/09,C12R1:91),(C12P5/00,C12R1:19),
PC (C12P5/00,C12R1:84),(C12P5/00,C12R1:91),(C12P5/00,C12R1:865),
PC (C12P5/00,C12R1:645),(C12P17/18,C12R1:645),(C12P17/18,C12R1:19) PC
PC (C12P17/18,C12R1:91),(C12P17/18,C12R1:84),(C12P17/18,C12R1:865) PC
C12N15/00
CC C12N15/00,C12N5/00,(C12N15/00,C12R1:91)
CC Start codon (ARG/Met) starts at nucleic acid position (6) and
CC stop codon
CC (TGA) at position 1644
PH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..2076
/organism="Artemisia annua"
/mol_type="genomic DNA"
/db_xref="taxon:35608"

FEATURES
source
Query Match 99.3%; Score 1637.4; DB 6; Length 2076;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

ORIGIN
QY 2 CATGGCCTTACAGAAACCACTTATCGCCCATTCGCCAATTCCTCCCAAGCATTTG 61
Db 5 CATGTCCTTACAGAAACCACTTATCGCCCATTCGCCAATTCCTCCCAAGCATTTG 64
QY 62 GGGAGATCAGTTTCTCATCTATCAAAACCAAGTAGACGAGCGGTGGCAACAGATAGTCAA 121

65 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGACGAGCGGTGGCAACAGATAGTCAA 124
QY 122 TGATTTAAAAAAGAGTGGGCAACTACTAAAGAAGCTTTGATATTTCCATGAACA 181
Db 125 TGATTTAAAAAAGAGTGGGCAACTACTAAAGAAGCTTTGATATTTCCATGAACA 184
QY 182 TGCCAAATTTGTTGAAGCTGATGATGAAATTCACAGCTTTGGAATACCGTATCACTTTGA 241
Db 185 TGCCAAATTTGTTGAAGCTGATGATGAAATTCACAGCTTTGGAATACCGTATCACTTTGA 244
QY 242 ACGGAGATGATCATGCAATTTGCAATGATTTATGAACAATATGCGTGAATACTGGAATGG 301
Db 245 ACGGAGATGATCATGCAATTTGCAATGATTTATGAACAATATGCGTGAATACTGGAATGG 304
QY 302 TGACCGCTCTTCTTATGCTTCCGCTCTTATGCGAAAGCAAGGATATTTATGTTACATGTA 361
Db 305 TGACCGCTCTTCTTATGCTTCCGCTCTTATGCGAAAGCAAGGATATTTATGTTACATGTA 364
QY 362 TGTTTTCAATACTATAAAGACAAATAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGT 421
Db 365 TGTTTTCAATACTATAAAGACAAATAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGT 424
QY 422 TGAAGGTTTCTGTTGATGTTGTAAGCAACTTCTATGAGGCTACCTGGGAGATATATTT 481
Db 425 TGAAGGTTTCTGTTGATGTTGTAAGCAACTTCTATGAGGCTACCTGGGAGATATATTT 484
QY 482 AGAAGATGCTCTTGGTTTTTACAGATCTCGTCTTAGCATTTAGCAAAAGATGCTTTTTC 541
Db 485 AGAAGATGCTCTTGGTTTTTACAGATCTCGTCTTAGCATTTAGCAAAAGATGCTTTTTC 544
QY 542 TACAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAG 601
Db 545 TACAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAG 604
QY 602 GTTGCAAGAATAGAGCGGCGCAGTACACTCTTCTTATCAACAACAAGATTCATATA 661
Db 605 GTTGCAAGAATAGAGCGGCGCAGTACACTCTTCTTATCAACAACAAGATTCATATA 664
QY 662 CAAGACTTTTACTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCATTGCAAGGA 721
Db 665 CAAGACTTTTACTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCATTGCAAGGA 724
QY 722 AGAGCTCAGCCATGTTGTGCAAAATGGTGGAAAGCTTTGCGATATCAAGAAGAACCGACCTTG 781
Db 725 AGAGCTCAGCCATGTTGTGCAAAATGGTGGAAAGCTTTGCGATATCAAGAAGAACCGACCTTG 784
QY 782 TTTAAGAGATAGAATTTGTAATGCTACTTTTGGGACTAGGTTCAAGCTATAGGCCACA 841
Db 785 TTTAAGAGATAGGTTTGTGTAATGCTACTTTTGGGACTAGGTTCAAGCTATAGGCCACA 844
QY 842 GTATTTCCCGGCTAGAGTTTCTTCAACAAGCTTTGCTGTTATTAATCTCTTATAGATGA 901
Db 845 GTATTTCCCGGCTAGAGTTTCTTCAACAAGCTTTGCTGTTATTAATCTCTTATAGATGA 904
QY 902 CACTTATGATGCTATGCTTATGAAGAACTTAAAGATCTTTACTGAAGCTGTTGAAAG 961
Db 905 CACTTATGATGCTATGCTTATGAAGAACTTAAAGATCTTTACTGAAGCTGTTGAAAG 964
QY 962 GTGGTCAATTTACATGCTTAGACACACTTCAGAAATACATGAACCGATATACAAATTTAT 1021
Db 965 GTGGTCAATTTACATGCTTAGACACACTTCAGAAATACATGAACCGATATACAAATTTAT 1024
QY 1022 CATGGATACATACAGAAATGGAAGAAATTTCTGCAAGAGGGAAGAACAGATCTATT 1081
Db 1025 CATGGATACATACAGAAATGGAAGAAATTTCTGCAAGAGGGAAGAACAGATCTATT 1084
QY 1082 TAACATGCGCAAGAAATTTGTTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATG 1141
Db 1085 TAACATGCGCAAGAAATTTGTTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATG 1144
QY 1142 GCGAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTCTGG 1201
Db 1145 GCGAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTCTGG 1204

```
QY 1202 CGGTGCTAACCTGCTTACACCACTTGTATCTTGGCATGAGTGATATATTCACAAAGA 1261
Db 1205 CGGTGCTAACCTGCTTACACCACTTGTATCTTGGCATGAGTGATATATTCACAAAGA 1264
QY 1262 GTCTGTCGAATGGGCTGCTCTGCACCTCTCTTTTAGTACTCTCAGGTATATCTTGTGTCG 1321
Db 1265 GTCTGTCGAATGGGCTGCTCTGCACCTCTCTTTTAGTACTCTCAGGTATATCTTGTGTCG 1324
QY 1322 AGCCCTAAATGATCTCATGACCCACACAGGCGGACGAGCAAGAAACATAGTTCATCGAG 1381
Db 1325 AGCCCTAAATGATCTCATGACCCACACAGGCGGACGAGCAAGAAACATAGTTCATCGAG 1384
QY 1382 CTTTGAAGCTTATATGAAGGATATATGTCATGAGGAGTATGCCCAACCTTGATTGA 1441
Db 1385 CTTTGAAGCTTATATGAAGGATATATGTCATGAGGAGTATGCCCAACCTTGATTGA 1444
QY 1442 CAAGGAAGTAGAAGATGTGTGGAAGATATAACCGAGAGTACTCTCAACATAAAACAT 1501
Db 1445 CAAGGAAGTAGAAGATGTGTGGAAGATATAACCGAGAGTACTCTCAACATAAAACAT 1504
QY 1502 TCCAGGCGGTTATTGATGGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCATATGC 1561
Db 1505 TCCAGGCGGTTATTGATGGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCATATGC 1564
QY 1562 AGMAAGGATTAATCTCACACGTATGGGAGACGATACAAACATCTCTATAAGTCTCTACT 1621
Db 1565 AGMAAGGATTAATCTCACACGTATGGGAGACGATACAAACATCTCTATAAGTCTCTACT 1624
QY 1622 CGTTATCTCTATGATATATGAGGATC 1648
Db 1625 CGTTATCTCTATGATATATGACTACC 1651

RESULT 3
LOCUS AF138959 2080 bp mRNA linear PLN 25-SEP-2000
DEFINITION Artemisia annua amorpho-4,11-diene synthase mRNA, complete cds.
ACCESSION AF138959
VERSION AF138959.1 GI:7381216
KEYWORDS
SOURCE
ORGANISM
  Artemisia annua (sweet wormwood)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; campanulids; Asterales; Asteraceae; Asteroideae;
  Anthemideae; Artemisia.
REFERENCE
  Mercke, P., Bengtsson, M., Bouwmeester, H.J., Posthumus, M.A. and
  Brodelius, P.E.
  Molecular cloning, expression, and characterization of
  amorpho-4,11-diene synthase, a key enzyme of artemisinin
  biosynthesis in Artemisia annua L
  Arch. Biochem. Biophys. 381 (2), 173-180 (2000)
  20485153
MEDLINE 11032404
PUBMED
REFERENCE
  2 (bases 1 to 2080)
  Mercke, P.E. and Brodelius, P.E.
  Direct Submission
  Submitted (26-MAR-1999) Plant Biochemistry, Lund University, P.O.
  Box 117, Lund S-221 00, Sweden
  Location/Qualifiers
    1..2080
      /organism="Artemisia annua"
      /mol_type="mRNA"
      /db_xref="taxon:35608"
      /tissue_type="leaf"
      /note="Artemisia annua L."
    13..1653
      /function="catalyzes the production of amorpho-4,11-diene
      from farnesyl diphosphate"
      /note="sesquiterpene synthase"
      /codon_start=1
  CDS
```

```
/product="amorpho-4,11-diene synthase"
/protein_id="AAF61439.1"
/db_xref="GI:7381217"
translation="MSLSEKPIRPIANPPSIWGDQFLIYKQVQEQVEIQVINDLKK
EVQLKALDIPMKHANLLKIDIEIQRLGIPYHFERIDHALQCIYETFGNWDGR
SSLMFLMRKQGYVTVCDVFNKYKNGAFQSLANDVEGLLELYEATSMRVPGEIIL
EDALGFTSRSLISIMTKDAFSPALFTBIQRLKQPLWKRPLRIEAAQVIFPYQODS
HNKTLKIALKLEFNLOSLEHLSHVCKWKAQDIKKNAPCLRDRIVCYIPFWGLSGS
YEPQYSRARVPFTKAVAVITLIDDTYDAYGVYEELKIPTAEVRSITCLDTLPEYMK
PIYKLFMTTYTEBEEFLAKGRTDLDFNCGKEFVKFVRNLVWEAKWANGHGLPTTEBH
DPVVIITGGANLTFTCYLGSMDITFKESVWASAPPLFRFVGLGRRLNDLMTTHA
EQERKSSSLESYKNYVNEEYATLIYKEVEDWKNIDINREYLTITKNIIPRLMLAV
IYLQCFLEVOVAGKDNFTRMGDEVKHLIKSLIVYPMI"
```

ORIGIN

```
Query Match 99.2%; Score 1635.8; DB 8; Length 2080;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CATGGCACTTACAGAGAAAACCTATTTCGCCCATTTGCCAACTTTCCTCCCAAGCAATTG 61
Db 12 CATGTCACTTACAGAGAAAACCTATTTCGCCCATTTGCCAACTTTCCTCCCAAGCAATTG 71
QY 62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAA CAGATAGTGA 121
Db 72 GGGAGATCAGTTTCTCATCTATGAAAAGCAAGTAGAGCAAGGGGTGGAA CAGATAGTGA 131
QY 122 TGATTTAAAAAAGAGTGGCGCACTACTATAAAGAGCTTTGGATATTCCTATGAACA 181
Db 132 TGATTTAAAAAAGAGTGGCGCACTACTATAAAGAGCTTTGGATATTCCTATGAACA 191
QY 182 TGCCAAATTTGTTGAAGCTGATTGATGAATCAACGCCCTTGGAAATACCGTATCACATTGA 241
Db 192 TGCCAAATTTGTTGAAGCTGATTGATGAATCAACGCCCTTGGAAATACCGTATCACATTGA 251
QY 242 ACGGGAGATTGATCATGCAATTCGAATGATTTATGAAACATATGTTGATGATAACTGGAATGG 301
Db 252 ACGGGAGATTGATCATGCAATTCGAATGATTTATGAAACATATGTTGATGATAACTGGAATGG 311
QY 302 TGACCGCTCTTCCTTATGTTCCGTCCTTATGCGAAGCAAGGATATATGTTACATGTGA 361
Db 312 TGACCGCTCTTCCTTATGTTCCGTCCTTATGCGAAGCAAGGATATATGTTACATGTGA 371
QY 362 TGTTTTCAATAACTATAAAGACAAAATAGGACGCTTCAAGCAATCGTTAGCTAATGATGT 421
Db 372 TGTTTTCAATAACTATAAAGACAAAATAGGACGCTTCAAGCAATCGTTAGCTAATGATGT 431
QY 422 TGAAGTTTGTGTGATGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATATT 481
Db 432 TGAAGTTTGTGTGATGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATATT 491
QY 482 AGRAGATCTCTTGGTTTTACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTTC 541
Db 492 AGRAGATCTCTTGGTTTTACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTTC 551
QY 542 TACAAACCCCGCTCTTTTATCCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAG 601
Db 552 TACAAACCCCGCTCTTTTATCCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAG 611
QY 602 GTTGCACAAAGATAGAGGGCGGCGAGTACATTCCTTTCTATCAACAACAAGATTCTCATAA 661
Db 612 GTTGCACAAAGATAGAGGGCGGCGAGTACATTCCTTTCTATCAACAACAAGATTCTCATAA 671
QY 662 CAAGACTTACTTAAACTGCTAAGTTAGAGTTCAATTTGCTTCAGTTCATTCGACAGGA 721
Db 672 CAAGACTTACTTAAACTGCTAAGTTAGAGTTCAATTTGCTTCAGTTCATTCGACAGGA 731
QY 722 AGAGCTCAGCCATGTGTCAAAATGTTGAAAGAGCTTTTCGATATCAAGAAGAACCCAGCTTG 781
Db 732 AGAGCTCAGCCATGTGTCAAAATGTTGAAAGAGCTTTTCGATATCAAGAAGAACCCAGCTTG 791
QY 782 TTAAAGAGATAGAATTGTTGAATGCTATTTTGGGAGTACGTTTGGGCTATGAGGCACA 841
```


Db 792 TTTAAGAGATAGAAATGTTGAATCTACTTTTGGGAGTAGGTCAGGCTATGAGCCACA 851
QY 842 GTATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTGCTGTATTAATCTTTATAGATGA 901
Db 852 GTATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTGCTGTATTAATCTTTATAGATGA 911
QY 902 CACTTATGATCGTATGCTATGATGAAGAACTTAAAGATCTTTAAGCTGTGGAAG 961
Db 912 CACTTATGATCGTATGCTATGATGAAGAACTTAAAGATCTTTAAGCTGTGGAAG 971
QY 962 GTGGTCAATTTACATGCTTTAGACACACTTCCAGATACATGAAACCGATATACAAATATT 1021
Db 972 GTGGTCAATTTACATGCTTTAGACACACTTCCAGATACATGAAACCGATATACAAATATT 1031
QY 1022 CATGGATACATACACAAATGGAAGAAATTTCTTCAAAAGGAGGAGAAACAGATCTATT 1081
Db 1032 CATGGATACATACACAAATGGAAGAAATTTCTTCAAAAGGAGGAGAAACAGATCTATT 1091
QY 1082 TAACTCGGCAAGAAATTTGGAAGAGTTTGTAGAAACCTGATGTTGAGGCAAAATG 1141
Db 1092 TAACTCGGCAAGAAATTTGGAAGAGTTTGTAGAAACCTGATGTTGAGGCAAAATG 1151
QY 1142 GGCAAATGAGGACACATACCAACCACTGGAAGCATGATCCAGTTGTAATCATCTG 1201
Db 1152 GGCAAATGAGGACACATACCAACCACTGGAAGCATGATCCAGTTGTAATCATCTG 1211
QY 1202 CGGTGCTAACTGCTTACAAACACTTGTATCTTTGGCATGATGATATATTCACAAAGA 1261
Db 1212 CGGTGCTAACTGCTTACAAACACTTGTATCTTTGGCATGATGATATATTCACAAAGA 1271
QY 1262 GTCTGCGAATGGGCTGCTCTGCACTCTCTCTTTTATAGATCTACGATATCTGCTG 1321
Db 1272 GTCTGCGAATGGGCTGCTCTGCACTCTCTCTTTTATAGATCTACGATATCTGCTG 1331
QY 1322 ACCCTTAATGATCTCATGACCCACAGGCGGCAAGAAAGAAACATGTTTCATCGAG 1381
Db 1332 ACCCTTAATGATCTCATGACCCACAGGCGGCAAGAAAGAAACATGTTTCATCGAG 1391
QY 1382 CTTGAAAGTTATATGAGAAATATATGTCAATGAGAGATGATGCCAAACCTTGATTTA 1441
Db 1392 CTTGAAAGTTATATGAGAAATATATGTCAATGAGAGATGATGCCAAACCTTGATTTA 1451
QY 1442 CAAGGAAGTATAGATGTGGAAAGATATATAACCGAGAGTACCTCAACATAAAACAT 1501
Db 1452 CAAGGAAGTATAGATGTGGAAAGATATATAACCGAGAGTACCTCAACATAAAACAT 1511
QY 1502 TCCAGGCGCTTATGATGCTGTGATCTATTGTCGCACTTTCTTGAAGTTCATATGC 1561
Db 1512 TCCAGGCGCTTATGATGCTGTGATCTATTGTCGCACTTTCTTGAAGTTCATATGC 1571
QY 1562 AGGAAAGGATAAATTACACGATGGGAGACGAATACAAACATCTCAAAAAGTCTCTACT 1621
Db 1572 AGGAAAGGATAAATTACACGATGGGAGACGAATACAAACATCTCAAAAAGTCTCTACT 1631
QY 1622 CGTTTATCCATGATGATATGAGGATC 1648
Db 1632 CGTTTATCCATGATGATATGAGGATC 1658

RESULT 4
AA251751
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AA251751 2106 bp mRNA linear PLN 24-NOV-2000
Artemisia annua mRNA for amorpha-4,11-diene synthase (kcs12 gene).
AJ251751
AJ251751.1 GI:8546577
amorpha-4,11-diene synthase; kcs12 gene.
Artemisia annua (sweet wormwood)
Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroidae;
Anthemideae; Artemisia.
1

AUTHORS Chang, Y.J., Song, S.H., Park, S.H. and Kim, S.U.
TITLE Amorpha-4,11-diene synthase of Artemisia annua: cDNA isolation and bacterial expression of a terpene synthase involved in artemisinin biosynthesis
JOURNAL Arch. Biochem. Biophys. 383 (2), 178-184 (2000)
MEDLINE 21036108
PUBMED 11185551
REFERENCE 2 (bases 1 to 2106)
AUTHORS Chang, Y.J.
TITLE Direct Submision
JOURNAL Submitted (10-DEC-1999) Chang Y.J., Research Center for New Bio-Materials in Agriculture, Seoul National University, Kwonsun-gu, Soedun-dong 103, Suwon 441-744, SOUTH KOREA
FEATURES
source
1. .2106
Location/Qualifiers
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/country="South Korea:Suwon"
1. .2106
/gene="kcs12"
34. .1675
/gene="kcs12"
/function="sesquiterpene cyclase"
/codon_start=1
/product="amorpha-4,11-diene synthase"
/protein_id="CA894691.1"
/db_xref="GI:8546578"
/db_xref="GOA:O9LW98"
/db_xref="SPTREMBL:O9LW98"
translation="MSLTDEKPIRPINPPSPISWGDQFLIYKQVEQGVQIIVNDLKK
EYRQLKALDIPMKHANLLKIDETIRGIPYHFERIDHALQCIYETVGNWNGDR
SSLWRLMKQGYVTVCDVNNYKDKNGAFKQSLANDVEGLLELYEATSRVPEIIL
EDALGFTSRSLSMTKDAPSTNPAFTQIRALKQPLWKLPRIEAAYIPFQQQDS
ENKTLKAKLEFNLLQSLHKEBELSHVCKWKAFLDKNAPCLDRDITVCCVFWGLSG
PEPQTSRAVFEFTKAVANTLIDDTYDAYGVYBELKFTFAVERWSITCLDTLPEYMK
FYIKDFMDTYTEMEFEKAGRTDLFNCCKEFVKEFVRNLNMAKWNANEHIFITEEH
DPVITLGGANLLTTCYLGMSDIFPKSEVWAVSAPPLPRYSGLGRRLNDLMTHTKA
EQERXHSNLLSEYMKVNEBYAQTLIYKEVDYMKDINREYLTNTKNTPRLLMAV
IYLCFLQVYAGKDNFTRMGDYKHLIKSLIVYPMISI"

Query Match 99.1%; Score 1634.2; DB 8; Length 2106;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1639; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CATGCACTTACAGAGAAAACCTATTGCGCCCATTTGCCAATTTCTCTCCACGATTG 61
Db 33 CATGCACTTACAGAGAAAACCTATTGCGCCCATTTGCCAATTTCTCTCCACGATTG 92
QY 62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGAGGGGTGGAACAGATAGTAA 121
Db 93 GGGAGATCAGTTTCTCATCTATGAAAGCAAGTAGAGAGGGGTGGAACAGATAGTAA 152
QY 122 TGATTTAAAAAGAGTGGGCACTACTATAAAGAGCTTTTGATATTCTCTATGAACA 181
Db 153 TGATTTAAAAAGAGTGGGCACTACTATAAAGAGCTTTTGATATTCTCTATGAACA 212
QY 182 TGCCCAATTTGTTGAAGCTGATTGATGAAATTCACCGCTTTGGAAATCCGTATCACTTTGA 241
Db 213 TGCCCAATTTGTTGAAGCTGATTGATGAAATTCACCGCTTTGGAAATCCGTATCACTTTGA 272
QY 242 ACGGAGATTTGATCATGCTGCAATTCATTTATGAAACAATATGGTGATAACTGGAATGG 301
Db 273 ACGGAGATTTGATCATGCTGCAATTCATTTATGAAACAATATGGTGATAACTGGAATGG 332
QY 302 TGACCGCTCTTCTTATGTTGCTTATGCGAAAGCAAGATATTTATGTTACATGTA 361
Db 333 TGACCGCTCTTCTTATGTTGCTTATGCGAAAGCAAGATATTTATGTTACATGTA 392
QY 362 TCTTTTCAATTAATAAGACAAATATGAGGCTTCAAGCAATCGTTAGCTATGATGT 421
Db 393 TGTTTTCAATTAATAAGACAAATATGAGGCTTCAAGCAATCGTTAGCTATGATGT 452

422 TGAAGGTTTGTGAGTGTGACGAAGCACTTCTATGAGGGTACCTGGGAGATTATATT 481
453 TGAAGGTTTGTGAGTGTGACGAAGCACTTCTATGAGGGTACCTGGGAGATTATATT 512
482 AGAAGATGCTTGTGTTTACACGATCTCGTCTTAGCATATGACAAAGATGCTTTTTC 541
513 AGAAGATGCTTGTGTTTACACGATCTCGTCTTAGCATATGACAAAGATGCTTTTTC 572
542 TACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAG 601
573 TACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAG 632
602 GTTGCCAAAGATAGAGCGCGGAGTACATCTCTTCTATCAACAAAGATCTCTATAA 661
633 GTTGCCAAAGATAGAGCGCGGAGTACATCTCTTCTATCAACAAAGATCTCTATAA 692
662 CAAGACTTTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCTCAGTCATTGACAAAG 721
693 CAAGACTTTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCTCAGTCATTGACAAAG 752
722 AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGACGACCTTG 781
753 AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGACGACCTTG 812
782 TTTAGAGATAGAAATTTGATGCTACTTTTGGGAGTACTAGTTTCAAGCTTATGAGCCACA 841
813 TTTAGAGATAGAAATTTGATGCTACTTTTGGGAGTACTAGTTTCAAGCTTATGAGCCACA 872
842 GTATTCCTCCGGCTAGAGTTTCTTCAAAAGCTTTTCTGTTTATCTTCTATAGATGA 901
873 GTATTCCTCCGGCTAGAGTTTCTTCAAAAGCTTTTCTGTTTATCTTCTATAGATGA 932
902 CACTTATGATGCTATGCTATGATGAAGAACTTAAAGATCTTCTAGTGAAGCTTGAAG 961
933 CACTTATGATGCTATGCTATGATGAAGAACTTAAAGATCTTCTAGTGAAGCTTGAAG 992
962 GTGCTCAATATACATGCTTAGACACACTTCCAGAAATACATGAACCCGATATACAAATATT 1021
993 GTGCTCAATATACATGCTTAGACACACTTCCAGAAATACATGAACCCGATATACAAATATT 1052
1022 CATGGATATACACAGAAATGGAAGATTTCTTCAAAAGCTTTTCTGTTTATCTTCTAGTGAAG 1081
1053 CATGGATATACACAGAAATGGAAGATTTCTTCAAAAGCTTTTCTGTTTATCTTCTAGTGAAG 1112
1082 TAACTCGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATG 1141
1113 TAACTCGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATG 1172
1142 GGCAATAGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGATATCAATCTG 1201
1173 GGCAATAGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGATATCAATCTG 1232
1202 CGGTGCTAACCTGCTTACACACACTTGTATCTTGGCATGAGTGATATTCACAAAGA 1261
1233 CGGTGCTAACCTGCTTACACACACTTGTATCTTGGCATGAGTGATATTCACAAAGA 1292
1262 GTCTGTCGAATGGCTGTCTCTCACCTCTCTTTTAGTACTACGATATCTTGTGTCG 1321
1293 GTCTGTCGAATGGCTGTCTCTCACCTCTCTTTTAGTACTACGATATCTTGTGTCG 1352
1322 AGCCTTAATGATCTCATGACCCACAGCGGACGAGCAAGAAAGAAACATAGTTCATCGAG 1381
1353 AGCCTTAATGATCTCATGACCCACAGCGGACGAGCAAGAAAGAAACATAGTTCATCGAG 1412
1382 CTTTGAAGTTATACAGGAATATATGTCATGAGGATGATGCCAAACCTTGATTTA 1441
1413 CTTTGAAGTTATACAGGAATATATGTCATGAGGATGATGCCAAACCTTGATTTA 1472
1442 CAAGGAAGTAGAAGATGTTGTGGAAGATATAAACCAGGAGTACCTCAACTAAACAT 1501
1473 CAAGGAAGTAGAAGATGTTGTGGAAGATATAAACCAGGAGTACCTCAACTAAACAT 1532

QY 1502 TCCAAAGCCGTTATGATGGCTGTGATCTATTTGTCGAGTTTCTTGAAGTTCAATATGC 1561
Db 1533 TCCAAAGCCGTTATGATGGCTGTGATCTATTTGTCGAGTTTCTTGAAGTTCAATATGC 1592
QY 1562 AGGAAAGGATAACTTTCACAGTATGGGAGACGAATACAAACATCTCTATAAGTCTTACT 1621
Db 1593 AGGAAAGGATAACTTTCACAGTATGGGAGACGAATACAAACATCTCTATAAGTCTTACT 1652
QY 1622 CGTTATCTCTATGATGATATATGAGGATC 1648
Db 1653 CGTTATCTCTATGATGATATATGAGTACC 1679

RESULT 5
AF327526
LOCUS Artemisia annua amorpho-4,11-diene synthase (AMS1) mRNA, complete cds.
DEFINITION AF327526 2067 bp mRNA linear PLN 05-MAR-2001
ACCESSION AF327526
VERSION AF327526.1 GI:13195431
KEYWORDS
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae; Artemisia.

REFERENCE 1 (bases 1 to 2067)
AUTHORS Liu, Y., Ye, H.C. and Li, G.F.
TITLE Cloning of sesquiterpene cyclase cDNA from Artemisia annua
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2067)
AUTHORS Liu, Y., Ye, H.C. and Li, G.F.
TITLE Direct Submission

JOURNAL
Submitted (08-DEC-2000) Department of Cell and Gene Engineering,
Institute of Botany, Chinese Academy of Sciences, Xiangshan,
Nanxincun 20, Beijing 100093, P.R. China
FEATURES
Location/Qualifiers
1..2067
/organism="Artemisia annua"
/mol_type="mRNA"
/strain="001"
/db_xref="taxon:35608"
/tissue_type="leaf"
/country="China: Sichuan"
1..2067
/gene="AMS1"
13..1653
/gene="AMS1"
/note="sesquiterpene cyclase"
/codon_start=1
/product="amorpho-4,11-diene synthase"
/protein_id="AAK15696.1"
/db_xref="GI:13195432"
/translation="MSLTETPEIRPIANFPSPISMGDQLIYEKQVEQVIVNDLKK
EVRLLEALDIPMKHANLLKLDIQRGLPIYHFEQEDHDLQCIYETGDNWDGR
SSLFRLRYKQGYVTCDFNNYKDKGAFKQSLANDVGLLEALVEATSRVGEIML
EDALGFRSRLS IMTKDAPSTNPAFTIQRALKQFLWKRLPIERAAQITPFYQODS
HNKTLKLEFNLSLHKEELSHVCKWKAFDIKKNAPCLRDRIVECYFWGLSGG
PEPOYSRAVFTKAVAVITLIDDTYDAGTVEELKIFTEAVERWSITCLDTPYMK
PIYKLPDWTTEMEFLAKGRTDLFNGCKPEKFEVNLVWEAKWANGHPIPTTEH
DPVITLGGANLLTTCYLGMSDIFTKESVWASVAPPFRYSGLGRINDLMTSKA
EQEKHSSTLESYMKENVEAYQTLIYKEVEDWVDKINREYLTTRKINPRLLMAV
IYLQCFLEVOYAGNFTNRMGDEYNHLIKLSLLVYPMIS"

FEATURES
source
1..2067
/organism="Artemisia annua"
/mol_type="mRNA"
/strain="001"
/db_xref="taxon:35608"
/tissue_type="leaf"
/country="China: Sichuan"
1..2067
/gene="AMS1"
13..1653
/gene="AMS1"
/note="sesquiterpene cyclase"
/codon_start=1
/product="amorpho-4,11-diene synthase"
/protein_id="AAK15696.1"
/db_xref="GI:13195432"
/translation="MSLTETPEIRPIANFPSPISMGDQLIYEKQVEQVIVNDLKK
EVRLLEALDIPMKHANLLKLDIQRGLPIYHFEQEDHDLQCIYETGDNWDGR
SSLFRLRYKQGYVTCDFNNYKDKGAFKQSLANDVGLLEALVEATSRVGEIML
EDALGFRSRLS IMTKDAPSTNPAFTIQRALKQFLWKRLPIERAAQITPFYQODS
HNKTLKLEFNLSLHKEELSHVCKWKAFDIKKNAPCLRDRIVECYFWGLSGG
PEPOYSRAVFTKAVAVITLIDDTYDAGTVEELKIFTEAVERWSITCLDTPYMK
PIYKLPDWTTEMEFLAKGRTDLFNGCKPEKFEVNLVWEAKWANGHPIPTTEH
DPVITLGGANLLTTCYLGMSDIFTKESVWASVAPPFRYSGLGRINDLMTSKA
EQEKHSSTLESYMKENVEAYQTLIYKEVEDWVDKINREYLTTRKINPRLLMAV
IYLQCFLEVOYAGNFTNRMGDEYNHLIKLSLLVYPMIS"

gene
CDS
cn: Farnesyl phosphate
pyrophosphatase
cyclase

Query Match 98.6%; Score 1626.2; DB 8; Length 2067;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1634; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 2 CATGGCACTTACAGCAAGAAACCTATTGCGCCCATTTGCCACTTTCCTCCAGCAATTG 61
Db 12 CATGTCACTTACAGCAAGAAACCTATTGCGCCCATTTGCCACTTTCCTCCAGCAATTG 71

ORIGIN
Query Match 98.6%; Score 1626.2; DB 8; Length 2067;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1634; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 2 CATGGCACTTACAGCAAGAAACCTATTGCGCCCATTTGCCACTTTCCTCCAGCAATTG 61
Db 12 CATGTCACTTACAGCAAGAAACCTATTGCGCCCATTTGCCACTTTCCTCCAGCAATTG 71

```
QY 62 GGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTAA 121
Db 72 GGAGATCAGTTTCTCATCTATGAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTAA 131
QY 122 TGATTTAAAAAGAGAGTGGCGCAACTACTAAAGAAAGCTTTGGATATTCCTATGAACA 181
Db 132 TGATTTAAAAAGAGAGTGGCGCAACTACTAAAGAAAGCTTTGGATATTCCTATGAACA 191
QY 182 TGCAATTTGTTGAAGCTGATGATGAATTCACAGCGCTTGGAAATACCGTATCACTTTGA 241
Db 192 TGCAATTTGTTGAAGCTGATGATGAATTCACAGCGCTTGGAAATACCGTATCACTTTGA 251
QY 242 ACCGGAGATTGATCATGCATTGCAATGATTTATGAAAACATATGGTGATTAACCTGGAATGG 301
Db 252 ACAGGAGATTGATCATGCATTGCAATGATTTATGAAAACATATGGTGATTAACCTGGAATGG 311
QY 302 TGACCGCTCTTCCTTATGTTTCCTGTTTATGCGAAAGCAAGATATTTATGTTACATGTA 361
Db 312 TGACCGCTCTTCCTTATGTTTCCTGTTTATGCGAAAGCAAGATATTTATGTTACATGTA 371
QY 362 TGTTTCAATAAATATTAAGCAAAATGAGCGTTCAAGCAATCGTTAGCTTAATGATGT 421
Db 372 TGTTTCAATAAATATTAAGCAAAATGAGCGTTCAAGCAATCGTTAGCTTAATGATGT 431
QY 422 TGAAGTTTGTCTGAGTTGTACGAAGCAACTTCTATGAGGTTACCTGGGAGATTAAT 481
Db 432 TGAAGTTTGTCTGAGTTGTACGAAGCAACTTCTATGAGGTTACCTGGGAGATTAAT 491
QY 482 AGAAGATGCTCTTGTTTTCACGATCTGCTTTAGCAATTTATGACAAAGATGCTTTTTC 541
Db 492 AGAAGATGCTCTTGTTTTCACGATCTGCTTTAGCAATTTATGACAAAGATGCTTTTTC 551
QY 542 TACAAACCCCGCTCTTTTACCGAAATACACGGGCCTTAAGCAACCCCTTTGAAAG 601
Db 552 TACAAACCCCGCTCTTTTACCGAAATACACGGGCCTTAAGCAACCCCTTTGAAAG 611
QY 602 GTTGCCAAAGATAGAGCGGCGAGTACATCTCTTTCTATCAACCAAGATTTCTATAA 661
Db 612 GTTGCCAAAGATAGAGCGGCGAGTACATCTCTTTCTATCAACCAAGATTTCTATAA 671
QY 662 CAAGACTTTACTTAACTTGTCTTAACTTGAAGTCAATTTGCTTCACTTGCATGCAAGGA 721
Db 672 CAAGACTTTACTTAACTTGTCTTAACTTGAAGTCAATTTGCTTCACTTGCATGCAAGGA 731
QY 722 AGAGCTAGCATGCTGCTCAATGCTGGAAGCTTTGATATCAAGAAAGCAAGCTTTG 781
Db 732 AGAGCTAGCATGCTGCTCAATGCTGGAAGCTTTGATATCAAGAAAGCAAGCTTTG 791
QY 782 TTTAAGATAGAAATGTTGATGCTACTTTTGGGACTAGGTTCAAGGCTATGAGCCACA 841
Db 792 TTTAAGATAGAAATGTTGATGCTACTTTTGGGACTAGGTTCAAGGCTTTGAGCCACA 851
QY 842 GATTTCCCGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTAACTCTTATAGATGA 901
Db 852 GATTTCCCGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTAACTCTTATAGATGA 911
QY 902 CACTTATGATGCTGATGTTGATGTTATGAAAGCTTAAAGCTTTTACTGAAAGCTTTGAAAG 961
Db 912 CACTTATGATGCTGATGTTGATGTTATGAAAGCTTAAAGCTTTTACTGAAAGCTTTGAAAG 971
QY 962 GTGGTCAATTACATGCTTACACACTTCCAGAAATACATGAAACCGATATACAAATTAAT 1021
Db 972 GTGGTCAATTACATGCTTACACACTTCCAGAAATACATGAAACCGATATACAAATTAAT 1031
QY 1022 CATGATATACATACAGAAATGGAAGAAATTTCTTCAAGAGGAGGGAAGAACAGATCTATT 1081
Db 1032 CATGATATACATACAGAAATGGAAGAAATTTCTTCAAGAGGAGGGAAGAACAGATCTATT 1091
QY 1082 TAACTGCGGCAAGAAATTTGTTGAAAGAGTTTGTAGAAACCTGATGTTGAAACCAATG 1141
Db 1092 TAACTGCGGCAAGAAATTTGTTGAAAGAGTTTGTAGAAACCTGATGTTGAAACCAATG 1151
```

```
QY 1142 GGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATCTATTACTGG 1201
Db 1152 GGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATCTATTACTGG 1211
QY 1202 CGTCTCTAACCTGCTTACAAACAACTTGTATCTTGGCATGAGTGATATATTCACAAAAGA 1261
Db 1212 CGTCTCTAACCTGCTTACAAACAACTTGTATCTTGGCATGAGTGATATATTCACAAAAGA 1271
QY 1262 GTCTGTGCAATCGGCTGCTCTGCACTCTCTCTTTTGTAGATCTCAGGTATCTTGGTGG 1321
Db 1272 GTCTGTGCAATCGGCTGCTCTGCACTCTCTCTTTTGTAGATCTCAGGTATCTTGGTGG 1331
QY 1322 AGCCCTAATGATCTCATGACCAACCAAGCGGAGCAAGAAAGAAACATAGTTTCATCGAG 1381
Db 1332 AGCCCTAATGATCTCATGACCAACCAAGCGGAGCAAGAAAGAAACATAGTTTCATCGAG 1391
QY 1382 CTTTGAAGCTTATATGAAGGAATATATGTCATGAGGAGTATGCCCAACCTTGATTTA 1441
Db 1392 CTTTGAAGCTTATATGAAGGAATATATGTCATGAGGAGTATGCCCAACCTTGATTTA 1451
QY 1442 CAAGGAAGTAGAAGTGTGTGGAAGATATAAACCAGAGTACCTCACTCACTAAACAT 1501
Db 1452 CAAGGAAGTAGAAGTGTGTGGAAGATATAAACCAGAGTACCTCACTCACTAAACAT 1511
QY 1502 TCCAGGCCGTTTATGATGCTGCTGATCTATTTGTGCGAGTTTCTTGAAGTTCAATATGC 1561
Db 1512 TCCAGGCCGTTTATGATGCTGCTGATCTATTTGTGCGAGTTTCTTGAAGTTCAATATGC 1571
QY 1562 AGGAAAGGTAACCTTCCACAGTATGGAGACGATACAAACATCTCTATAAAGTCTCTACT 1621
Db 1572 AGGAAAGGTAACCTTCCACAGTATGGAGACGATACAAACATCTCTATAAAGTCTCTACT 1631
QY 1622 CGTTTATCTCTATGATATATGAGGATC 1648
Db 1632 CGTTTATCTCTATGATATATGACTACC 1658
```

```
RESULT 6
AF472361 1902 bp mRNA linear PIN 03-JAN-2003
LOCUS Artemisia annua beta-caryophyllene synthase QHS1 (QHS1) mRNA,
DEFINITION complete cds.
ACCESSION AF472361
VERSION AF472361.1 GI:18766933
```

```
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
REFERENCE 1 (bases 1 to 1902)
AUTHORS Cai, Y., Jia, J.-W., Crock, J., Lin, Z.-X., Chen, X.-Y. and Croteau, R.
TITLE A cDNA clone for beta-caryophyllene synthase from Artemisia annua
JOURNAL Phytochemistry 61 (5), 523-529 (2002)
MEDLINE 22297925
PUBMED 12409018
```

```
REFERENCE 2 (bases 1 to 1902)
AUTHORS Cai, Y., Crock, J., Jia, J.-W., Lin, Z.-X., Croteau, R. and Chen, X.-Y.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-2002) National Laboratory of Plant Molecular
Genetics, Institute of Plant Physiology and Ecology, Shanghai
Institutes for Biological Sciences, Chinese Academy of Sciences,
300 Fenglin Road, Shanghai 200032, P. R. China
```

```
FEATURES
source Location/Qualifiers
  ..1902
  /organism="Artemisia annua"
  /mol_type="mRNA"
  /db_xref="taxon:35608"
  ..1902
  /gene="QHS1"
  24..1670
  /gene="QHS1"
```

/note="sesquiterpene cyclase"
/codon_start=1
/product="beta-caryophyllene synthase QHS1"
/protein_id="AAL79181.1"
/db_xref="gi:18766934"
/translation="MSVKEEKVIRPIVHPHPSVWADQFLIFDDKQAEQANQVNVNEL
REVDKLVSLDVOTETNLILKIDAIQRLGIAVHPPEEIEQALQHIYDVTGDDWK
RSPSLWPIILROQGFYVSCDIFKNYKXSDGSKESLNDVDELLELYEATYLRVQEG
VLDALVPTRCLEKIDLVHNTPLSYIOEALQKPUHKLRLTEARURYIPNTEQG
ASHNESLKLAKLFLNKLQSLKESLSEVSRWMLGDPVNNLPYARDRVECYFWALG
YFFPKYSQARIFLAKVLSLATVLDYDAYGTYEELKIFTEAIQRWSITCDMLPEY
LKLIVQGVLDVYIEMEETWEGKAHLISYAKESKPEFIRSYMEAKWANEQVPTFA
EHMSVAFVSSGYMLATTCFVGMGDIVTDEAPKWLTKPPIIKASCATARLMDDTHSQ
KEEKERLHVASVSYSKQYDVTEEHVLKVENKJEDAKDITRESLVKDIKPMPLMM
RVINLAQWMDVLYKHKGDTWVGEELKDHIKSLLVHPPIPI"

ORIGIN

Query Match 52.0%; Score 857; DB 8; Length 1902;
Best Local Similarity 70.8%; Pred. No. 1.4e-186;
Matches 1168; Conservative 0; Mismatches 475; Indels 6; Gaps 2;
QY 2 CATGCCATTACAGAGAAAACCTATTTCGCCCATTTGCGCACTTTCTCCCAAGCATTTTG 61
DB 23 CATGCTCTTTAAGAGAGAAAGTAAATTCGGGCCCATTTGCTCCACTTTCTCTCCAGCGTTTG 82
QY 62 GGGAGATCAGTTCTCTCATCT---ATCAAAGCAAGTAGAGCAAG---GGGTGGACAGAT 115
DB 83 GCGACAGATTTCTCATCTTCGATGACAGCAAGCAAGCAAGCAAGCAAGTGTAGAACAAAGT 142
QY 116 AGTGAATGATTTAAAAAAGAGTGGCGCAACTACTTAAAGAGGCTTTGGATATTCCTAT 175
DB 143 AGTCACAGTTAAGAGAGATGTGAGAAAGATTTAGTGTCTATCTTTAGATGTTCAAAC 202
QY 176 GAACATGCCAAATTTGTTGAGCTGATTTGAATGAATGAATCAACGCCCTTGGAAATCCGGTATCA 235
DB 203 GGAGCACACAAATTTGCTAAACTTATTGATGCTATCCAAACGCTTTGGTATAGCGGTATCA 262
QY 236 CTTTGAACGGGAGATGATCATGCAATGCAATGTTATGAAACATATGTTGATGATGATG 295
DB 263 TTTTGAAGAGAGATGTAGCAAGCCTTACACATATTTATGATACATATGTTGATGATG 322
QY 296 GAATGGTGACCGCTCTCCCTTATGTTCCGCTTATGCGAAAGCAAGGATATTTATGTTAC 355
DB 323 GAAAGGCAGAGCCCTTCCCTTTGTTTTCGAATCCTTCGGCAACAGGTTTTTATGTTTC 382
QY 356 ATGTGATGTTTTCAATACTATAAGACAAATAAGAGCGGTTCAAGCAATCGTTAGCTAA 415
DB 383 TTGTGATATTTTAAAAACTATAAAAAAGAGATGGTTCAATTTAAGGAATCCCTCACCAA 442
QY 416 TGATGTTGAAGGTTGCTTGTAGTCTGACGAACAACTTCTATGAGGTTACCTGGGAGAT 475
DB 443 CGATGTAAGGCTTGTCTGAGCTGTATGAGCGACATATTTGAGAGTCAAGGGGAAG 502
QY 476 TATATTAGAAGATGCTCTTGGTTTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGC 535
DB 503 GGTCTAGATGATGCTCTTGTCTTTCACAGGACTTGTCTTGAGAAAATAGCAAGATCT 562
QY 536 TTTTCTTCAAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTG 595
DB 563 TGTTCACACCAACCAACTATCTACCTATACATAAAGAGCACTAAAAACAGCCGTTTACA 622
QY 596 GAAAGGTTGCCAAGATAGAGGGCGGCGCAGTACATTCCTTCTATCAACACAGATTC 655
DB 623 TAAAGGTTGACAGACTAGAGGCAATGCGTTTACATTCCTATGTAGCAACACAGCATC 682
QY 656 TCATAACAAGACTTTTACTTAACTTGTAAAGTTAGAGTTTCAATTTGCTTTCAGTCAATTGCA 715
DB 683 TCATAATGAGTCTTTACTTCAACTTGGAGGCTAGGTTTCACTTACTTCAATCATTTGCA 742
QY 716 CAAGGAAGAGCTCAGCCATGTGTGCAATGTGGAAAGCTTTCCGATATCAAGAAAGACGC 775
DB 743 TAGAAAGAGGCTTAGCGAAGTTTCCAGGTTGTGGAAAGGCTTTGATGTGCCCAACAACTCT 802

QY 776 ACCTTGTTTTAAGAGATAGAAATTTGTAATGCTACTTTTGGGAGCTAGGTTCCAGGCTATCA 835
DB 803 ACCTTATGCAAGAGATAGAAATGTTGTAATGTTATTTTGGCAGCTAGGTTGCTATTTTGA 862
QY 836 GCCACAGTATTCOCGGGCTAGAGTTTCTTCCACAAAAGCTGTTGCTGTATATAACTCTTAT 895
DB 863 GCCAAAATATTCACAGCTAGGATCTTTTAGCAAAAAGTAATTTCCGCTAGCAACTGTTCT 922
QY 896 AGATGACACATTATGATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATG 955
DB 923 TGACGACACATTATGATGCTTTATGAAACCTATGAAAGAACTTTAAGATCTTTTACTGAAAGCAAT 982
QY 956 TGAAGGTGCTCAATTTACATGCTTTAGACACACATTTCCAGAAATACATGAAACCGATATACAA 1015
DB 983 TCAAGGTGCTCAATTTACATGCTTTAGACACACATTTCCAGAAATACCTGAAACTATTTATACA 1042
QY 1016 ATTATTCATGATATACATACACAGAAATGGAAGAAATTTCTTCAAGAGGAGGAGAAACAGA 1075
DB 1043 AGGAGTCTTGATATATACATAGAAATGGAAGAAATTAATGGAAGAGGAGGAGAAACGACA 1102
QY 1076 TCTATTTAATCTGGCGCAAGAAATTTGTAAGAGTGTGTTAGAAACCTGATGTTGTAAGC 1135
DB 1103 TCATCTTAGCTATGCTTAAAGAAATCTATGAAAGAGTGTATTTAGAAAGCTATATGATGGAAGC 1162
QY 1136 ARAATGGGCAATGAGGGACACATACCAACCACTGAAAGAGCATGATCCAGTTGTATATCAT 1195
DB 1163 ARAATGGGCAATGAGGGGTATGTACCAACAGCAGAGAGCATATGTCAGTTGCGTTGT 1222
QY 1196 TACTGGCGGTGCTAACTGCTTACCAACAACTGTTTATCTTGGCATGAGTGTATATTCAC 1255
DB 1223 AAGCAGCGGTATAGCATGCTTGCACCAACATGTTTGTGGCATGGGTATATCGTTAC 1282
QY 1256 ARAAGATGCTGCGAATGGGCTGCTCTGACCTCTCTTTTAGATACTCAGGTATATCT 1315
DB 1283 AGATGAAGCATTCAAATGGGCTCTGACAAAACCTCTCATCAAAAGCTTCATGTGCAAT 1342
QY 1316 TGTGTACGCGCTAAATGATCTCATGACCCCAAGCGCGAGCAAGAAAGAAACATAGTTC 1375
DB 1343 TGCTAGACTCATGATGATATCTACTCTCAAGAGGAGAGAAAGAAATACATGTTGC 1402
QY 1376 ATCGAGCCTTTGAAGTTATATGAAGGAATATAATGTCAATGAGAGTATGCCAAACCTT 1435
DB 1403 ATCTAGTGTGTTGAAAGTTTACATGAAGCAATATGATGTGACAGAGGAGCATGTCTTAAAGT 1462
QY 1436 GATTACAGGAAGTAGAAGATGTGTGAAAGATATATAACCGAGAGTACCTCACAACATA 1495
DB 1463 ATTAAACAGAAAATCGAGGATCGTGGAAAGATATAACCCGAGATCCCTTGTGCGTAA 1522
QY 1496 AAACATTTCCAAAGCCGTTATTGATGGCTGTGATCTATTTGGCCAGTTCCTTGAAGTTCA 1555
DB 1523 AGATTTCCGATGCTCTGATGATGCGAGTGAATTAACCTTGCCACAGGTGATGATGTTT 1582
QY 1556 ATATGCAAGGAAGGATAAATTCACGATGCGGAGACGAATACAAACATCTCATAAAGTC 1615
DB 1583 ATATAACATAAAGACGTTTTCATATGTGGGAGAAAGAACTCAAGGATCATATTTAAATC 1642
QY 1616 TCTACTGTTTATCTTATGATGATATGAG 1644
DB 1643 TTTGCTCGTTTCATCTTATACCTATATAAG 1671

RESULT 7
AAN271793
LOCUS
DEFINITION
Artemisia annua mRNA for putative sesquiterpene cyclase (cASC34 gene).
ACCESSION
AJ271793
VERSION
GI:10696989
KEYWORDS
cASC34 gene; sesquiterpene cyclase.
SOURCE
Artemisia annua
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

AAN271793 2026 bp mRNA linear PLN 04-OCT-2000
Artemisia annua mRNA for putative sesquiterpene cyclase (cASC34 gene).

asterids; campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae; Artemisia.

1 Van Geldre, E., De Pauw, I., Inze, D., Van Montagu, M. and Van den Beekhout, E.
Cloning and molecular analysis of two new sesquiterpene cyclases from *Artemisia annua* L.
Plant Sci. 158 (1-2), 163-171 (2000)
10996256
2 (bases 1 to 2026)
Van Geldre, E.G.L.
Direct Submission
Submitted (31-JAN-2000) Van Geldre E.G.L., Laboratory for Pharmaceutical Biotechnology, University of Ghent, Harelbekestraat 72, 9000 Gent, BELGIUM
Location/Qualifiers
1. .2026
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/tissue_type="leaves"
1. .2026
/gene="cASC34"
39. .1688
/gene="cASC34"
/codon_start=1
/evidence=experimental
/product="putative sesquiterpene cyclase"
/protein_id="CAC12732.1"
/db_xref="GI:10696990"
/db_xref="GOA:Q9FYX6"
/translation="MSVKEKVIPIVHPFPPSVADQFLIFDDKQAOANVEQVNL
RDVRDLVSLDQVQEHNLKLIADIAORLGHAFPEEIEQALQHIYTYGDWNG
RPSLWFLRLQGGFVVSQDIKPKYKEDGSIILRNPLTNDVEGLLELYEALYLRVOE
GWDDALVTRCLEKIALDVLHNTPLSTYIQEALQKPLQKLEALHLYIPMVEQ
QASHESLLKALFNLSLHKLNEVSRKGLDVPNNLPYARDRWVCFWAL
GVYFEPKYSQARIFLAKVLSLATVLDYDAYGTBELKIFTSIAIOWASTICIDMPE
YLKLYQGVLDIYIEEMIEGKGAHLISYAKESMKBFIRSYNMEAKWANEGVPTA
EHRMSVAFVSSGVSLATTCFVGNGDIVTDEAFKWLTKPPIIKASCAIARLMDIHS
OKEBERIHVASVESVMQYDVTBEHLVKFNKLEDAWKDITRESLVRKDIPLPLM
MRVINLAQVNDVLKHKDGFNVGBELKHLSLLVHPPI"

ORIGIN

Query Match 50.7%; Score 836; DB 8; Length 2026;
Best Local Similarity 70.4%; Pred. No. 9.5e-182;
Matches 1163; Conservative 0; Mismatches 480; Indels 9; Gaps 3;

QY 2 CATGGCACTACAGAGAAACCTATTCCGCCCAATGGCCAACTTTCCTCCAGCATTTG 61
DB 38 CATGCTCTGTTAAAGAGAGAAAGTAATTCGGCCCAATGTCCACTTTCTCTCCTAGCGTTG 97

QY 62 GGGAGATCAGTTTCTCATCT--ATCAAGAGCAAGTAGAGCAAG--GGGTGGACAGAT 115
DB 98 GGCAGACCAAGTTTCTCATCTTCGATGACAGCAAGCAGAGCAAGCAATGTAGAACAGT 157

QY 116 AGTGAATGATTTAAAAAGAGAGTGGCGCAACTACTAAAGAAAGCTTTGGATATTCCTAT 175
DB 158 AGTCAACGAGTTAAGAGAGATGTGAGAAAGATTAGTGTCACTTTAGATGTTCAAC 217

QY 176 GAAACATGCCAATTTGTGAAGCTGATGTGAGAAATTCACGCCCTTGGAAATACCGTATCA 235
DB 218 GGAGCACACAAATTTGTGTAAGCTTATGATGCTATCCAAAGCTTGGTATAGCGTATCA 277

QY 236 CTTTGAACGGGAGATGATCATGCAATGCAATGATTTATGAAACATATGGTGATACTG 295
DB 278 TTTTGAAGAGGAGATTGAGCAAGCCCTTACACATATTATGATACATATGTTGATGACTG 337

QY 296 GAATGTGACCGCTTCTCTTATGTTGGTCCGCTTCTATCGAAAGCAAGGATATTATGTTAC 355
DB 338 GAAAGGAGAGCCCTTCCCTTTGGTTGGATTCGTAATCCTTCGGCAACAAGGTTTATGTTTC 397

QY 356 ATGTGATGTTTCAATAACTATATAAGACAAAATGGAGC-----GTTCAAGCAATCGTTAGC 412

Db 398 TTGTGATATTTTAAAAAACHTATAAAAAAGAGATGTTCCATTTTAAAGAAATCCCTCAC 457

QY 413 TAATGATGTTGAAGGTTTGTGTTGAGTTGTAGCAAGCAACTTCTATGAGGGTACCTGGGA 472

DB 458 CAACGATGTAGAGGCTTGTGTTGAGCTGTATGAGGCGACATATTTGAGAGTGAAGCGGA 517

QY 473 GATTATATTAGAAGATGCTCTTGGTTTACACGATCTCGTCTTAGCATTTATGACAAAGA 532

DB 518 AGGGTTTCTAGATGATGCTCTTGGTTTTCACAGGAGCTTGTCTTGGAAATATGACAAAGA 577

QY 533 TGTCTTTTCTACAAACCCCGCTCTTTTACCGGAAATACAAACGGGCACTAAAGCAACCCCT 592

DB 578 TCTTGTTCACACCAACCCACACATATCTACCTACATACAAAGAGCACTAAACAGCGGT 637

QY 593 TTGAAAAGGTTGCCAAGATAGAGGCGGCGAGTACATTCCTTTCTATCAACAACAAGA 652

DB 638 AAAAAAAGGTTGGCAAGACTAGAGGCATTCATTACATTTCTATGTACGAACAACAAGC 697

QY 653 TTCTCATTAACAAGACTTTTACTTTAACTTGTCTTAAGTTTAGAGTTCAATTTGCTTCAGTCATT 712

DB 698 ATCTCACAAATGAGTCTTTACTCAAACTTGGGAAGCTAGGTTTCAACTTACTTCAATCATTT 757

QY 713 GCACAAGGAAGAGCTCAGCCATGTGTCAAAATGGTGGAAAGCTTTTCGATATCAAGAAGAA 772

DB 758 GCATAGAAAGAGGCTTAACGAAGTTTCCAGGTGGTGGAAAGGCTTGTATGTCCCAACAA 817

QY 773 CGCACCTTTTAAAGATAGAAATTTGTTGAATGCTACTTTTGGGAGCTAGGTTACGGCTA 832

DB 818 TCTACCTTATCAAGAGATAGAAATGTTGAATGTTATTTTGGGCACTAGGTTCTCTATTT 877

QY 833 TGAGCCACAGTATTTCCCGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTAACTCT 892

DB 878 TGAGCCAAAATATTCTCAAGCTAGGATCTTTTACAAAAGTAATTTCCGTAGCAACTGT 937

QY 893 TATAGATCACACTTATGATCGTATGCTATGCTATGATGAAGAACTTAAGATCTTTACTGAAGC 952

DB 938 TCTTGAACACACTTATGATGCTTATGGAACCTTATGAAGAACTTAAGATCTTTACTGAAGC 997

QY 953 TGTGAAAAGGTTGCTCAATTTACATGCTTTAGACACACTTCCAGAAATACATGAACCGATATA 1012

DB 998 AATTCAAAAGTGTGCGATTACATGATAGATGTTCTCCAGAATACCTGAAACTATTATATA 1057

QY 1013 CAATATTTCATGATATACATACACAGAAATGAAGAATTTCTTGAAGAGGAGGAAGAAC 1072

DB 1058 CCAAGGAGTCTTGGATATATATACATAGAATGAAGAATAATGGGAAAAGAGGGAAGAAGC 1117

QY 1073 AGATCTATTAACTCGCGCAAGAAATTTGTGAAAGAGTTTGTAGAAAACCTGATGTTGA 1132

DB 1118 ACATCATCTTAGCTATGCTAAAGATCTATGAAGAGTTTATTAGAAGCTATATATGATGA 1177

QY 1133 AGCAAAATGGGCAAAATGAGGGAACATACCAACCACTGAAGAGCATGATCCAGTTGTAAT 1192

DB 1178 AGCAAAATGGGCAAAATGAGGGAATGTACCAACAGCAGAGAGGCATATGTGATGCGTT 1237

QY 1193 CATTAAGCGGCTTAACTGCTTTACACCAACTTGTATCTTGGCATGATGATATATT 1252

DB 1238 TGTAAAGCGGCTTAGCATGCTTTGCCACACATGTTTGTGGCATGATGATTCGT 1297

QY 1253 CACAAAAGAGTCTGCGAATGGGCTGCTCTGCACCTCTCTTTTAGATACTCAGGTAT 1312

DB 1298 TACAGATGAAGCATTTCAAATGGGCTCTGACAAAACCTCTATCATCAAAGCTTCATGTC 1357

QY 1313 ACTTGGTGCAGCCCTAAATGATCTCATGACCCCAAGGCGGAGCAAGAAAGAAACATAG 1372

DB 1358 AATTCTGACTCATGATGATATTCACTCTCAAAAGGAGGAGAAAGAAAGAAATACATGT 1417

QY 1373 TTCTCGAGCCCTTGAAGTTTATAGGAATATATGTCAATGAGGAGTATGCCCAAC 1432

DB 1418 TGCACTCTAGTGTGAAGTTTACATGAAGCAATATGATGTGACAGAGGAGCATGTCTCTTAA 1477

QY 1433 CTTGATTTTACAGGAGTAGAAGATGTGTGAAAGATATATAAAGAGAGTACCTCAAC 1492

Db 1478 AGTATTTAACAGAAAAATCAGGATCGGTGGAAAGATATATACCCGAGAAATCCCTTGTGCG 1537

QY 1493 TAAAAACATCCAGCGCGTATTATGATGGCTGTGATCTATTATTTGGCAGATTCTTGAAGT 1552

Db 1538 TAAAGATATTCAGATGCTCTGATGATGCGAGTGATTAACCTGGCAGCGATGGATGT 1597

QY 1553 TCAATATGACGAGAAAGGATAACCTTACACGCTATGGGAGAGAAATACAAACATCTCATAAA 1612

Db 1598 TTTATATAACATAAGACCGGTTTCACTAATGCGGAGAGAACTCAAGATCATATATA 1657

QY 1613 GTCTCTACTGTTTATCTCTAATGATATATAGAG 1644

Db 1658 ATCTTTGCTCGTTTCATCTCTATACCTATATATAAG 1689

RESULT 8

SCA304452

LOCUS

DEFINITION Solidago canadensis mRNA for germacrene A synthase (1 gene).

ACCESSION AJ304452

VERSION AJ304452.1 GI:13624777

KEYWORDS 1 gene; germacrene A synthase.

SOURCE Solidago canadensis

ORGANISM Solidago canadensis

1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Campanulids; Asterales; Asteraceae; Asteroideae; Asteraceae; Solidago.

1 Prosser, I., Phillips, A.L., Gittings, S., Lewis, M.J., Hooper, A.M., Pickett, J.A. and Beale, M.H.

(+)-(10R)-Germacrene A synthase from goldenrod, Solidago canadensis; cDNA isolation, bacterial expression and functional analysis

JOURNAL Phytochemistry 60 (7), 691-702 (2002)

MEDLINE 22126967

PUBMED 12127586

REFERENCE 2 (bases 1 to 1976)

AUTHORS Prosser, I.M.

DIRECT SUBMISSION

TITLE Submitted (11-DEC-2000) Prosser I.M., Plant Sciences, IACR-Long Ashton Research Station, Long Ashton, BS41 9AF, UNITED KINGDOM

JOURNAL Location/Qualifiers

FEATURES

source

1. .1976

/organism="Solidago canadensis"

/mol_type="mRNA"

/cultivar="Gold Baby"

/db_xref="taxon:59297"

/clone="SC1"

/tissue_type="leaf"

1. .1976

/gene="1"

41. .1690

/gene="1"

/function="conversion of farnesyl diphosphate (FPP) to (+)-germacrene A"

/codon_start=1

/evidence=experimental

/product="germacrene A synthase"

/protein_id="CAC36896.1"

/db_xref="GI:13624778"

/db_xref="GOA:Q9A67"

/db_xref="SPTREML:Q9A67"

/translation="MAAKQVEIRPVANVHPSLWGQFLHYDQEDHEVVDQIELL KEETREILIASDDPKYHNLLKLDIVQRLGIAYFHEHITQALDHIYSYVDGWNIGRTSLWFLRLQQGVYSCDIFNIYKLDNGSPKSLTKOIECMLELYEAYMRVQGEI ILDEALEFTKLEHIAKDPKRCNLTLSRIHEALRFPVKRLPRLDALRYIPFEQQ DSHNKLRLAKLGNRLQSLKSKLWKFDPKRLPVDRDLVELYFWILG VYFEQYSRSLFLTKTIRMAALLDDYDIYGYEELEFTKAVQKWSLTCMTLDY MKVIYKSLDLDVYEMEEIEIKQKAYQVHYAKESMIDLVTSTMYEARLWHEHVDPD EHNVSNTIYGGYKLTASSFVGMHGDIVTQESFKWLNPNPLIKASD.SRIMNDIVG HKEEQKHIAKSVENMYKYNLAEDVDYFLKERVEDAKWDKINRETLTKDKIHMALK MPPINLARVMDMLYKNGDNLKNVQGEIQDYMKSCFINPMSV"

ORIGIN

Query Match 47.11%; Score 776.2; DB 8; Length 1976;
Best Local Similarity 67.8%; Pred. No. 5.5e-168;
Matches 1119; Conservative 0; Mismatches 523; Indels 9; Gaps 2;

QY 2 CATGCACTTACAGAGAAACACCTATTGCGCCCATTTGCCAACTTTCCTCCAAGCATTTG 61

Db 40 CATGCTCTTAAACAAGTAGAGGTATTGCGCCAGTTGCAAACTATCATCCAAGCCTTTG 99

QY 62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCA-----AGGGTGGAAACAGAT 115

Db 100 GGGAGATCAGTTTCTACACTATGATGAGCAAGAGGATGAGCGTTGAGGTAGATCAACA 159

QY 116 AGTGAATGATTTAAAAAAGAGTGGCGCAACTACTAAAGAGCTTTTGGATATTCCTAT 175

Db 160 AATCGAAATTTTGAAGGAGAAACCGGAAAGAAATACTTTGCAAGTTTGGATGATCCAAC 219

QY 176 GAAACATGCCAATTTGTTGAAGCTGATGAAATTTCAACGCTTTGGAATACCGTATCA 235

Db 220 AAACATACAAATTTGCTGAAGCTGATGATGTTTATCAACGCTCTCGGTATAGCCTATTA 279

QY 236 CTTTGAACGGGAGATTGATCATGCAATGATGTTATTTATGAACAATATGGTGAATAGTG 295

Db 280 TTTTGAACATGAGATTACGCAAGCATTCGACCATATTTATAGTGTATACGGTGATGAATG 339

QY 296 GAATGCTGACCGCTTTCTCTTATGTTGCTTCTTATGCGAAAGCAAGGATATTTATCTAC 355

Db 340 GAATGCTGCTCGTACTTCCCTTTGCTTCTCGGCTCTCCGACCAACAGGGCTTTTACGTTTC 399

QY 356 ATGCGATCTTTTCAATAACTATAAAGCAAAATGGAGCGTTCAAGCAATCGTTAGCTAA 415

Db 400 ATGATATATTTTCAATATCTACAGCTTGATATGATGATCTTTCAAGGATTCCTTAACCA 459

QY 416 TGATGTTGAAGGTTTGTGTTGATGTTGACGAAGCAACTCTTATGAGGGTACCTGGGGAGAT 475

Db 460 GGTATTTGATGATGCTGCTTGTGTTTATGAGGAGCGCTATATGAGGTTGCAAGCGGAAAT 519

QY 476 TATATTAGAAGATGCTCTTGTGTTTATGACGATCTGCTTACCATTTATGACAAAGATGC 535

Db 520 CATTTAGATGAGGCTCTTGTGTTTATGAGTTTACAAAACACTCATCTTGAACACATTTGCAAGGATCC 579

QY 536 TTTTCTTACAAACCCCGCTCTTTTACCGAAATACAAAGGCGCACTTAAAGCAACCCCTTTG 595

Db 580 ACTTGTGTGCAACACACGCTCTCTAGACACATACATGAGCACTAGAGGGGCTGTACA 639

QY 596 GAAAGGTTGCCAAGATAGAGGGCGGCGAGPACATTTCTTTCTATCAACAACAAGATTC 655

Db 640 GAAAGGTTGCCAAGACTAGATGCAATACGATACATACCTTTCTATGAACAACAAGATTC 699

QY 656 TCATAACAAGACTTTTACTTAACTTGTAACTGTAAGTAGAGTTCAATTTGCTTCACTGATGCA 715

Db 700 TCACAACAAGTCTTACTTAAGCTTGCAAAGTGGGGTTCACCCGGCTTCAATCTCTGCA 759

QY 716 CAAGGAGAGCTCAGCCCATGTTGCAATGTTGGAAGCTTTTCGATATCAAGAAGAACGC 775

Db 760 TAAAGAGGCTTAGCCCACTTTTCCAAATGTTGGAAGAAATTTGATGCTCCAAGATCT 819

QY 776 ACCTTGTTTAAGAGATGAATTTGTTGAATGCTACTTTTGGGAGCTAGGTTTCAGGCTATGA 835

Db 820 ACCTTACGTAAGAGATAGATTGGTTGAACTCTACTTTTGGATACCTAGGTGTCTACTTCGA 879

QY 836 GCCACAGTATTCGGGGCTAGAGTTTCTTACAAAAGCTGTTGCTGTTTAACTCTTAT 895

Db 880 ACCTCAATATTTCTCGTTCGAGATATTTCTTGACAAAAACAATTAATAATGGCAGCAATCT 939

QY 896 AGATGACACTTATGATGGTATGTTTATGTAAGAACTTAAGATCTTTTACTTGAAGCTGT 955

Db 940 AGACGACAGTATGATATCTATGTTACTTACGAAGACTTGAATATTCACCAAGCCGT 999

QY 956 TGAAGGTGGTCAATTAATCATCTCTTAGACACACTTCCAGATACATGAACCAAGATACAA 1015

Db 1000 TCAAGGTTGGTCAATTAATCATCTCTTAGACACACTTCCAGATTAATGATGATTTATTA 1059


```
Qy 1016 ATTATTCATGATACATACACAGAAATGGAAGATTTCTTGCAAGGAGGAGAACAGA 1075
Db 1060 GAGCCTCTTGATGTTTATGAAGAAATGGAAGAAATCATAGAAAGGATGGAAGACATA 1119
Qy 1076 TCTATTTAACTGCGGCAAGAAATTTGTGAAGAGTTGTTAGAAACCTGATGTTGAAGC 1135
Db 1120 TCAAGTTCACATATGCAAGAGAGTCAATGATAGATTGGTTACAGTATATGACCAAGC 1179
Qy 1136 AAAATGGCAATATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTGTGAATCAT 1195
Db 1180 AAAATGGTTATCATGAGGGTCATGTGCCAATTTGACGAGCATAACTCAGTTACAAACAT 1239
Qy 1196 TACTGGCGGTGCTTAACCTGCTTACACAACTTTGTTATCTTGGCATG -- AGTGATATATT 1252
Db 1240 AACTGGTGGCTATAAAATGCTTACAGCATCAAGCTTTGTTGGCATGCAATGGTATATAGT 1299
Qy 1253 CACAAAAGAGTCTGTGCAATGGGCTGTCTCTGCACCTCTCTTTTATAGATACTCAGGTAT 1312
Db 1300 TACACAGAGCTTTCAATGGGTTCTCAACAACTCTCCACTTATAAAGCTTCATCTGA 1359
Qy 1313 ACTTGGTCGAGCCCTTAATGATCTCATGATGCCACCAAGGCCGAGCAAGAAAGAAACATAG 1372
Db 1360 CATTAGTAGGATTATGAATGATATCGTCCGTACAGGAGGAGCAACAAAGAAAGCATAT 1419
Qy 1373 TTCATCGAGCCTTGAAGTTATATGAAGAAATATATGTCATGAGGAGTATGCCCAAC 1432
Db 1420 TGCACTAGTGTGGAATATGATGAAGAAATATATCTCGCGGAGGAGCACTCATGA 1479
Qy 1433 CTGTGATTACAAAGAAAGTAGAAGATGTGTGGAAGATATAAACCAGAGAGTACCTCAAC 1492
Db 1480 TTTTCTCAAGAAAGAGTTGAAGATGATGGAAGATATAACCGAGAAACATTAACATG 1539
Qy 1493 TAAAAACATTCGAAGCCGTTATGATGCTGTGATCTATTTGTGCGCATGTTCTTGAAGT 1552
Db 1540 TAAAGACATTCATATGGCTCTTAAAGATGCTCCGATCAACCTGGCAGCGGTAATGATAT 1599
Qy 1553 TCAATATCGAGGAAGGATACTTACACGATGAGGAGACGATACAAACATCTCATAA 1612
Db 1600 GCTATACAAAACCGTGATATTTTAAAAAATGTGGACAAGAAATCCAGATTTATATGA 1659
Qy 1613 GTCTCTACTCGTTTATCTCATGATATATGA 1643
Db 1660 ATCTGTTTCATTAATCTATGAGTGTTCGA 1690

RESULT 9
LOCUS ATERPENE 1946 bp mRNA linear PLN 15-SEP-2000
DEFINITION Artemisia annua mRNA for epi-cedrol synthase.
ACCESSION AJ001539
VERSION AJ001539.1 GI:10183622
KEYWORDS epi-cedrol synthase.
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
1
REFERENCE
AUTHORS Mercke,P., Crock,J., Croteau,R. and Brodelius,P.E.
TITLE Cloning, expression, and characterization of epi-cedrol synthase, a
sequiterpene cyclase from Artemisia annua L
JOURNAL Arch. Biochem. Biophys. 369 (2), 213-222 (1999)
MEDLINE 99417501
PUBMED 10486140
REFERENCE
AUTHORS Brodelius,P.E.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1997) Brodelius P.E., Department of Plant
Biochemistry, Lund University, P.O.Box 117, Lund, S-22100, SWEDEN
FEATURES
Location/Qualifiers
1..1946
source
```

```
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/tissue_type="leaf"
7..10
TATA_signal
CDS
34..1677
/codon_start=1
/product="epi-cedrol synthase"
/protein_id="CAC08805.1"
/db_xref="GI:10183623"
/db_xref="GOA:Q9LLR9"
/db_xref="SPTREMBL:Q9LLR9"
/translation="MSLIVEDVIRPNANFPSEIWDQFLAYDQDQEQVEQVVKDLKE
EVSELLTALNSPTQHTTELLKFIDAIRLGATVFEERINOVFOKMYTAYGDKWTGKN
TSIWFRLMRHGFEVSDSDISTYKDKGRFKESLEKDVHGLELYEAAVMFVPGEGIL
DVALFRTCLDEATKPSLNSAVSSIREALTQPLJHKLPRLEALRVIPIYQQOAS
HSTLLKXALGFNLQSLHKLKLSIIKSWKSFVANNLPFARNRPECVYFWALAVY
FEQYSSSRVFLSRFFSIQTFLDDTYDAYTTELEQFTEAIOQRSITCLDGLPESMK
LIFQMLVKIPEEIEIILSKDKQHVNYIKETLKAQVQSYMTAEARKEEYIPTIEEH
TKYSYLSIGYKLALVAGFACMGDVIADDSFEMVFNPLVNACCLLCRMDDLGSHKG
EQBRKHVASITIECMKQFDASEQQAQYVESLKKKVEDAKKEINREFMITCKDVNIHVAMR
VLNFRSVDVLYIKKHDFHTVGVVEVINHIKSLFVDALIT"
1929
polyA_site
ORIGIN
Query Match 43.9%; Score 723.4; DB 8; Length 1946;
Best Local Similarity 65.4%; Pred. No. 7.7e-156;
Matches 1076; Conservative 0; Mismatches 566; Indels 3; Gaps 1;
Qy 2 CATGGCACTTACAGAAGAAACCACTATTTCGCCCATTCGCCCACTTCCCAACTTTCTCCCAAGCATTTG 61
Db 33 CATGCTCTTATAGTAGAGGATGTCATACGCCCAATGCAAAATTTCTTCTGAAATTTG 92
Qy 62 GGGAGATCAGTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGACACATAGTGAA 121
Db 93 GGGAGATCAGTTCTTCGCTATGACCAAGATGAGCAAGAGGAGTTGAAACAAGTAAATCAA 152
Qy 122 TGATTTAAAAAGAAAGTGGCGCAACTACTAAAAAGACCTTTGGATATTCCTATGAAAA 181
Db 153 AGATTTGAAAGAAAGATGAAAGCGGAACCTATTGACAGCTCTAAATTTCTCCACTCAGCA 212
Qy 182 TGCCAAATTTGTTGAAGCTGATTTGATAAATTCACAGCCTTTGGAATACCGTATCACTTTGA 241
Db 213 TACGGAGTTGTTGAAATTTATTGACGCAATCAAGCGCTTGGTATTGGTATTATTTTGA 272
Qy 242 ACGGGAGATGATCATGATGCAATGATTTATTGAAACATATGATGATGATGATGATGATG 301
Db 273 AGAGGAGATCAACCAAGTATTTCCAACACATGATATGCTGATATGCTGATGATGATGATG 332
Qy 302 TGACCGCTCTTCCTTATGTTTCCGCTCTTATCGAAAGCAAGGATATTATGTTACATGTA 361
Db 333 TGGTAATACTTCCCTTTGGTTTCGACTCATGCGACACATGATGATTTTTCGTTTCAAGCGA 392
Qy 362 TGTTTTCAATAACTATAAAGACAAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGT 421
Db 393 TATTTTGTAGTACCTACAAAGACAAAGAGGAGCGTTTAAAGGAGTCTTTAGAAAAGGACGT 452
Qy 422 TGAAGTTTGGTTGAGTTGTACGAGCAACTCTATGAGGGTACCTCGGGGAGATTAATTT 481
Db 453 TCACGGGTACTTTCAGTTGTTATGAGCGGCATATATGTTGTCCTGGGGAAGGGATACT 512
Qy 482 AGAGATGCTCTTGGTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTTC 541
Db 513 AGATGATGCTCTTGTTTTCAAGAACTTGTTCTGATGAAATAGCGAAAATTCCTAGTCT 572
Qy 542 TACAAACCCCGCTCTTTTTCACGAAATACAAAGGCGCACTTAAAGCAACCCCTTTGAAAAAG 601
Db 573 AAGTAATCTGCTGCTTTCTTCCCAATACGTGAAGCTCTAACGCAACCTTTTACATAAAG 632
Qy 602 GTTCCCAAGAAATAGAGCGGCGCAGTACATTCCTTCTTATCAACACACAGATTCATATA 661
Db 633 ATTACCGAGACTAGAGGCGTTGCGCTACATACCTTTCTACCAACACACAGCTTCTCACAG 692
```


662 CAAGACTTTTACTTAACTTGGTAAAGTTAGAGTTCAATTTGCTTCACTGTCAGTCACAGGA 721
 693 TGAGACATTTTAAACTGGCCAAATTTAGGTTTCAACCACTTCAATCACTTCCAGAA 752
 722 AGAGCTACGCCATGTGTGCAAAATGGTGAAGCTTTTGCATATCAAGAAGAACCCCTTG 781
 753 AGAGCTTAGCATAATTTCAAGTGGTGAAGAAAGTTTTCGATGTTGCAACAATCTACCGTA 812
 782 TTTAGAGATAGAAATTTGTAAGTCTACTTTGGGACTAGGTTCAAGCTATGAGCACA 841
 813 TGCAGAAATATAGACCGGTTGAATGCTACTTTGGGCAATTAGCAGTACTTCCAGCTCA 872
 842 GTATTCCCGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTATTAACCTTATAGATGA 901
 873 ATATTCTGATCAAGGCTTCTTATCAAGATTTTTCATACAAACATTTCTTATGA 932
 902 CACTTATGATGCTGATGTTTATGAGAACTTAAGATCTTTTACTGAAGCTTTGAAAG 961
 933 CACTTATGATGCTGATGTTTATGAGAACTTGAACCAATTTTACTGAAGCAATTTCAAAG 992
 962 GTGGTCAATTACATGCTTACACACTTCCAGATACATGAAACCGATATACAAATATT 1021
 993 GTGGTCAATTACATGCTTCCGAGGCTTCTTCCGAAAGCATGAGCTTATATCCAAATGCT 1052
 1022 CATGATATACATACAGAAATGGAAGATTTCTTCAAAAGGAGGGAAGACAGATCTATT 1081
 1053 AGTGAAGATTTTGAAGAAATAGAGGAATATTGTCGAGGATGGAACAGCATCATGT 1112
 1082 TAACTGGGCAAGAAATTTGTGAAGAGTTGTTAGAAACCTGATGTTGAAACCAAAATG 1141
 1113 CAATACATCAAGAAACACTGAAGAGGCTGTCCTCAAGCTATACAGCAAGCAAGATG 1172
 1142 GCGAATGAGGACATACACCACTGACAGGATGATGATGATGATGATGATGATGATGATG 1201
 1173 GCGAAGAAAGAGATATATACCAAGATGAGGAGGACACAAAGATGTCATATATAGCAT 1232
 1202 CGGTGCTAACCTGCTTACAAACACTTGTATCTGCTGATGATGATGATGATGATGATGATG 1261
 1233 CGGCTATAACTTGGCTAGTAGCAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
 1262 GTCTGTCGAATGGCTGCTCTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1321
 1293 CTCAATTGAATGGGTTTCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1352
 1322 AGCCCTTAATGATCTCATGACCCACAGGCGGAGCAAGAAAGAAATAGTATCATCGAG 1381
 1353 AACCATGGATGATCTTGGTTCCTAAGGCGGAGCAAGATAGAAAGCATGTTGCCCTCCAC 1412
 1382 CCGTGAAGGTATATGAAGCAATATATGTCATGAGGATGATGATGATGATGATGATGATGATG 1441
 1413 TATTGATCTCATGAGCAATTTGATGCTAGCAGCAAGCAAGCTTACGAATCACTCAA 1472
 1442 CAAGGAAGTAGAAGATGTGTGGAAGATATAAACCAGAGTAC--CTCAAACTAAAAA 1498
 1473 CAAGAAAGTTGAAGATGATGCAAGAAAGAAATAAATAGAGATTCATGATCAGGTGAAGGA 1532
 1499 CATTCGAAGCCGTTATGATGCTGTGATCTATTTGTCGACGCTTCTTGAAGTTCAATA 1558
 1533 TGTTAATATTCATGATGATGATGAGGCTCTTAATTTTTCAGTTCGCTGATGCTCTTTA 1592
 1559 TGCAGAAAGGATACTTCAACGATGAGGAGAGCAATACAAACATCTCATAAAGTCTCT 1618
 1593 TAAGATAAGGATCATTTTACATGTTGAGGATGAGGCTGATTAATCATATCAATCTCT 1652
 1619 ACTCGTTTATCTTATGATATATGA 1643
 1653 TTTGTTGATGCCATCATATATGA 1677

RESULT 10
 AF157059 1985 bp mRNA linear PLN 01-MAY-2001
 LOCUS
 DEFINITION Artemisia annua 8-epicidrol synthase (Ecs1) mRNA, complete cds.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES
 source

gene
 CDS

AF157059
 AF157059.1 GI:8885993
 Artemisia annua (sweet wormwood)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Anthemideae; Artemisia.
 1 (bases 1 to 1985)
 Hua, L. and Matsuda, S.P.
 The molecular cloning of 8-epicidrol synthase from Artemisia annua
 Arch. Biochem. Biophys. 369 (2), 208-212 (1999)
 99417500
 10486139
 2 (bases 1 to 1985)
 Hua, L. and Matsuda, S.P.
 Direct Submission
 Submitted (06-JUN-1999) Chemistry Department, Rice University, 6100
 S. Main St., Houston, TX 77005, USA
 Location/Qualifiers
 1..1985
 /organism="Artemisia annua"
 /mol_type="mRNA"
 /db_xref="taxon:35608"
 1..1985
 /gene="Ecs1"
 42..1685
 /gene="Ecs1"
 /note="sesquiterpene synthase"
 /codon_start=1
 /product="8-epicidrol synthase"
 /protein_id="AAF80333.1"
 /db_xref="GI:8885994"
 /translation="MSLVEDVIRPNANFSEIKGQFLAVDDQSQEGVEQVVKLUKE
 EVSELLTALNSPTQHTTELLKIDAIERLGIAYFEEDINQVFOHYTAYGDKWTGNN
 TSWFRLMRQHGFEVSDIFSTYKDEGRFESLEKDVHGLLELYEALRYIPYQOQAS
 DDLVPTCTCLDEIAKPNLSNSAVSOIRALTOPLKRLPRLEALRYIPYQOQAS
 HSETLXLKLGELQSLKKELSIISKWKSQDVANLVPYARNRVECVFWALAVY
 FEPQSSRFLSRFFSIQFLDDTDYDAYGYEELQTEALQIRNSICLGLPESMK
 LIQMWLKIIEEILSKDGQKHVNYIKETLKEAVOSYMEARWAKEEYIPTLEEH
 TKVYSISIGIKLAVGACMGDVADDSFEWFTNPPLVNACCLLCTMDLGGSHKG
 EQQRKHVASTIECKQFDSQOAYSLNKKVEDAWKEINREFMITCKDVNIHVMR
 VLNPSRSDVLYKNKDHFTHVGEVINHKSLFVDAIIT"

ORIGIN

Query Match 43.9%; Score 723.4; DB 8; Length 1985;
 Best Local Similarity 65.4%; Pred. No. 7.7e-156;
 Matches 1076; Conservative 0; Mismatches 566; Indels 3; Gaps 1;

Qy 2 CATGGCACTTACAGAGAAACCTATTTCGCCCACTTCCCACTTCTCCCAAGCATTTG 61
 Db 41 CATGCTCTTATAGTAGAGGATGTCATACGCCCAATGCAATTTTCTTCTGAAATTTG 100
 Qy 62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGCAAGGGGTGGAAACAGATAGTGA 121
 Db 101 GGGAGATCAGTTTCTTTCCTATGACGATGAGCAAGGGAGTTGAACAAGTATCAA 160
 Qy 122 TGATTTAAAAAGAGTGGCGCAACTACTAAAGAGCTTTGGATATTTCTCTATGAACA 181
 Db 161 AGATTTGAAGAGAGTGAAGAGCGAACTATTGACAGCTCTAAATTTCTCCGACTCAGCA 220
 Qy 182 TGCCAATTTGTTGAGCTGATGATGAATTCACGCTTGGAAATACCGTATCCTTGA 241
 Db 221 TACGGAGTTGTGAAATTTATTGAGCAATCAGCGCTTGGTATTGCGTATTAATTTGA 280
 Qy 242 ACGGAGATTGATCATGCTTGCATTTGATTTATGAACAATATGTTGATGAACCTGGAATGG 301
 Db 281 AGAGGAGATCAACCAAGTATTTCCACACATGTATCTGATATGTTGATGAGTGAAGCAGG 340
 Qy 302 TGACCGCTTCTCTTATAGTTCCGCTCTTATGCGAAAGCAAGGATATTTATGTTACATGTA 361
 Db 341 TGGTAATACTTCCCTTTTGGTTTCGACTCATGCGAACAATGATTTTTCGTTTCAAGCGCA 400

QY	362	TCGTTTTCAATTA	CTATAAAGACAAAA	TGGAGCGTTCAAGCAATCGTTAGCTAAATGATGT	421
DB	401	TAATTTTTAGTAC	CTTCAAAAGACAAAGAGGACGTTTTAAGGAGCTTTT	TAGAAAGGACGT	460
QY	422	TGAAGGTTTGCT	TGAGTTGTACGAAGCAACTCTATGAGGGTACCTG	GGGAGATTATATT	481
DB	461	TCACGGGTTACT	TGAGTTGTATGAGCGGCATATATGTTGTGCTCGCTGGG	AAGGATACT	520
QY	482	AGAAAGTCTCT	TGTTTTTACAGCATCTGCTTAGCATATGACAAAGATGCTTTTT	CTTC	541
DB	521	AGATGATGCTCT	TGTTTTTACAGGAACCTTGCTGATGAATAGCGAAAAAT	CCCTAGTCT	580
QY	542	TACAAACCCCGCT	CTTTTTTACCGAAATACAACGGGCACCTAAAGCAACCCCT	TGAAAAAG	601
DB	581	AAGTAACCTGCT	GTCTTCCCAATACGTGAAGCTCTAACGCAACCTTTACAT	AAAAAG	640
QY	602	GTTGCCAAGAA	TAGAGCGGCGGAGTCAATTCCTTCTATCAACAACAAGATCT	CTCATAA	661
DB	641	ATTACCGGAGT	AGGCGCTGCGGTACATACCTTTCTACCAACAACAGCTTCT	CACAG	700
QY	662	CAAGACTTTACT	TTAAACTTGCTTAAGTTAGAGTTCAATTTGCTTCAGTCA	ATTGCACAGGA	721
DB	701	TGAGACATTTG	TAAACTCGGCCAAATTAGGTTCAACCAACTTCAATCACTT	CAACAAGAA	760
QY	722	AGAGCTCAGCCAT	TGTGCAAAATGTTGGAAGCTTTTCGATATCAAGAAGAA	CGCACTTG	781
DB	761	AGAGCTTAGCAT	AAATTTCCAAAGTGTGGAAGAAAGTTTCGATGTTCGAAA	CAATCTACCGTA	820
QY	782	TTTAGAGATAG	NAATTTGTAATGCTACTTTTGGGACTAGCTTCAGGCTAT	TAGGCCACA	841
DB	821	TGCAAGAAATAG	ACCGGTGGAATGCTACTTTTGGGCATTAGCAGTGTACT	TCGAGGCTCA	880
QY	842	GTATTCGCCGGT	TAGAGTTTTCTTCCAAAAAGCTGTGCTTTATAACTCTTAT	AGATGA	901
DB	881	ATATTCGTAAT	CAAGAGTCTTCTTATCAAGATTTTTTCAATACAAACATTT	CTTGATGA	940
QY	902	CACCTTAGAT	TGCGTATGTGAAGAACTTTAAGATCTTTACGAAGCTGTT	GAAAG	961
DB	941	CACCTTAGAT	TGCGTACTTACGAAGAACTTGCAAAATTTACTGAAGCAAT	TCAAAG	1000
QY	962	GTGCTCAAT	TACATGCTTAGACACACTTCCAGAATACATGAACCGATAT	ACAAATAT	1021
DB	1001	GTGGTCAAT	TAACATGCTTGAGCGGCTTCTGGAAGCATGAAGCTTAT	TTCAAATGCT	1060
QY	1022	CATGGAATAC	ATACACAGAAATGGAAGATTTCTTGCAAGGGGGAAGACAT	CTATT	1081
DB	1061	AGTGAAGAT	TTTTTGAAGAAATAGAGGAATATTTGTCGAAGGATGGA	AAACAGCATCATGT	1120
QY	1082	TAACTCGGG	CAAAAGAAATTTGTGAAGAGTTGTAGAAACCTGATGGT	TGAAGCAAAATG	1141
DB	1121	CAACTACAT	CAAAAGAAACACTGAAAGAGGCTGTCCAAAGCTACAT	CAGAAAGCAAGTG	1180
QY	1142	GGCAAAATG	AGGGACATACCAACCACTGAAGAGCATGATCCAGTTGT	GTAAATCATTTACTGG	1201
DB	1181	GGCAAAAGAG	ATATACCAACGATTGAGGAGCACAAAGTGTCTATAT	TATTAGCAT	1240
QY	1202	CGGTGCTAAC	CTGCTTACAAACAATTTGTTATCTTGCGCATGAGTGAT	ATATTTCAAAAAGA	1261
DB	1241	CGGCTATAA	ACTTTGCGCTAGTACAGGTTTGTCTGCTGGTGGTAT	TATTCGATGA	1300
QY	1262	GTCTGCGAAT	GGGCTGTCTGCACTCTCTCTTTTATAGTACTCAGG	TATACCTGGTGG	1321
DB	1301	CTCATTTGA	ATGGGTTTTTCCAAATCCCTCTCTTGTCAAATGCT	TGTGTCTACTTTGTAG	1360
QY	1322	ACGGCTAAAT	GATCTCATGCCCAAGGCGGAGCAAGAAAGAAACAT	AGTTTCATCGAG	1381
DB	1361	AACCATGAT	GAATCTTGGTTTTCCCATAGGGGAGCAAGATAGA	AAACATGTTCCCTCCAC	1420
QY	1382	CGTTGAAG	TTATAGAAGGAATATAATGTCAATGAGGAGTAT	GCCCAACCTTGATTTA	1441
DB	1421	TATTTGAAT	GCTCATAGAAGCAATTTGATGTAGGAGCAACAAG	CGCTACGAATCACTCAA	1480

QY	1442	CAAGGAAGTGAAGATGTGTGGAAGATATAAACCCAGAGTAC---CTCACAACTAAAAA	1498
Db	1481	CAAGAAAGTTGAAGATGTCATGGAAGAAATAATAGAGAGTTTCATGATCACGTGTGAAGA	1540
QY	1499	CATTCCAGGCCGTTATTGATGGCTGTGATCTATTTTGTCAGTTCTTGAAGTTCCTT	1558
Db	1541	TGTTAATATTCAATGATGATGAGGTCCTTAATTTTTCACGTTGGTGGATGTCCTTTA	1600
QY	1559	TGCAGGAAGGATAACTTTCACACGTCATGGGACGACGATAACAACATCTCTATAAGTCTCT	1618
Db	1601	TAAAGTAATAGGATCACTTTACACATGTTGGAGTAGAGGTGATAATCATATCAATCAATCTCT	1660
QY	1619	ACTCGTTTATCTATGAGTATAGA	1643
Db	1661	TTTCGTTGATGCCATCAATACATGA	1685

AF327527	4392 bp	DNA	linear	PLN 05-MAR-2001
Artemisia annua amorpha-4,11-diene synthase (AMS1) gene, complete cds				
AF327527	GI:13195433			
Artemisia annua (sweet wormwood)				
Artemisia annua				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae; Artemisia.				
1 (bases 1 to 4392)				
Liu, Y., Ye, H.C. and Li, G.F.				
Cloning of sesquiterpene cyclase gene from Artemisia annua				
Unpublished				
2 (bases 1 to 4392)				
Liu, Y., Ye, H.C. and Li, G.F.				
Direct Submission				
Submitted (08-DEC-2000) Department of Cell and Gene Engineering, Institute of Botany, Chinese Academy of Sciences, Xiangshan, Nanxincun 20, Beijing 100093, P.R. China				
Location/Qualifiers				
1. .4392				
/organism="Artemisia annua"				
/mol_type="genomic DNA"				
/strain="001"				
/db_xref="taxon:35608"				
/tissue_type="leaf"				
/country="China: Sichuan"				
<13. .>3978				
/gene="AMS1"				
join<13. .>99,214. .484,2183. .2564,2741. .2959,3088. .3226				
3224. .3572,3685. .>3978)				
/genes="AMS1"				
/product="amorpha-4,11-diene synthase"				
join(13. .>99,214. .484,2183. .2564,2741. .2959,3088. .3226				
3224. .3572,3685. .>3978)				
/gene="AMS1"				
/notes="sesquiterpene cyclase"				
/codon_start=1				
/product="amorpha-4,11-diene synthase"				
/protein_id="AAK15697.1"				
/db_xref="GI:13195434"				
/translation="MSLTBEKPIRPIANTPPSWGQFLIYEKQVEQVEQIVNDLKK EYVQLLKEALDIPMGHANLLKIDIEORLIGIPYHFEQIDHALQCIYETIGDNWDGR SLWFLMKRQGYVYTCDFNNYKDGAFKQSLANDVGLLELYEATSRVPEIGML EDALGTRSLRLSMTKDAFTSNPALFETQALQPLWKLPRIEAAQVTPFYQOQDS HNKTLILAKLENLQSLHKEELSHVCKWMAFDIKQAPCLRDRIVECYFVGLGSG PEQYGRARVFTKAVATLDDTVDYACTVEELKFTFAVERWSITCLDTPYMK PIYLFMDITNEEFLEKGRDGLNCSEKFEVKEFVRLMWEAKWANGSHIPTTSEH DFVITLITGGANLLTTCIGMSDITFKESVEWAVSAPPLFRYSGILGRRUNDLWTHA EQBRKTHSSLSIESMYKENVNIEYAQTLIYKEVEDVYKQILKRRINDLWTHA IYLCQFLEVOYAGKQNTBMGDYKHLISLYVPM51"				

RESULT 11	AF327527
LOCUS	Artemisia annua amorpho-4,11-diene synthase (AMS1) gene, complete cds
DEFINITION	Artemisia annua amorpho-4,11-diene synthase (AMS1) gene, complete cds
ACCESSION	AF327527
VERSION	AF327527.1
KEYWORDS	GI:13195433
SOURCE	Artemisia annua (sweet wormwood)
ORGANISM	Artemisia annua
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae; Artemisia.
AUTHORS	Liu, Y., Ye, H.C. and Li, G.F.
TITLE	Cloning of sesquiterpene cyclase gene from Artemisia annua
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 4392)
AUTHORS	Liu, Y., Ye, H.C. and Li, G.F.
TITLE	Direct Submission
JOURNAL	Submitted (08-DEC-2000) Department of Cell and Gene Engineering, Institute of Botany, Chinese Academy of Sciences, Xiangshan, Nanxincun 20, Beijing 100093, P.R. China
FEATURES	Location/Qualifiers
source	1. .4392
	/organism="Artemisia annua"
	/mol_type="genomic DNA"
	/strain="001"
	/db_xref="taxon:35608"
gene	/tissue_type="leaf"
	/country="China: Sichuan"
	<13. .>3978
	/gene="AMS1"
mRNA	join<13. .>99,214. .484,2183. .2564,2741. .2959,3088. .3226
	3224. .3572,3685. .>3978)
	/genes="AMS1"
	/product="amorpha-4,11-diene synthase"
	join(13. .>99,214. .484,2183. .2564,2741. .2959,3088. .3226
	3224. .3572,3685. .>3978)
	/gene="AMS1"
	/notes="sesquiterpene cyclase"
	/codon_start=1
	/product="amorpha-4,11-diene synthase"
	/protein_id="AAK15697.1"
	/db_xref="GI:13195434"
	/translation="MSLTBEKPIRPIANTPPSWGQFLIYEKQVEQVEQIVNDLKK EYVQLLKEALDIPMGHANLLKIDIEORLIGIPYHFEQIDHALQCIYETIGDNWDGR SLWFLMKRQGYVYTCDFNNYKDGAFKQSLANDVGLLELYEATSRVPEIGML EDALGTRSLRLSMTKDAFTSNPALFETQALQPLWKLPRIEAAQVTPFYQOQDS HNKTLILAKLENLQSLHKEELSHVCKWMAFDIKQAPCLRDRIVECYFVGLGSG PEQYGRARVFTKAVATLDDTVDYACTVEELKFTFAVERWSITCLDTPYMK PIYLFMDITNEEFLEKGRDGLNCSEKFEVKEFVRLMWEAKWANGSHIPTTSEH DFVITLITGGANLLTTCIGMSDITFKESVEWAVSAPPLFRYSGILGRRUNDLWTHA EQBRKTHSSLSIESMYKENVNIEYAQTLIYKEVEDVYKQILKRRINDLWTHA IYLCQFLEVOYAGKQNTBMGDYKHLISLYVPM51"

ORIGIN

Query Match 43.9%; Score 723.2; DB 8; Length 4392;
Best Local Similarity 71.1%; Pred. No. 8.1e-156;
Matches 1281; Conservative 0; Mismatches 8; Indels 513; Gaps 4;

QY	360	GATGTTTCAATACTATAAAGACAAAATGAGCGTTCACGCAATCGTTAGCTAATGAT	419
DB	2182	GATGTTTCAATACTATAAAGACAAAATGAGCGTTCACGCAATCGTTAGCTAATGAT	2241
QY	420	GTTGAAGGTTTGGTTGAGTTGACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATA	479
DB	2242	GTTGAAGGTTTGGTTGAGTTGACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATG	2301
QY	480	TTAGAAGAGTCTCTGGTTTACACGATCTGCTTAGCAATATGACAAAAGATGCTTTT	539
DB	2302	TTAGAAGAGTCTCTGGTTTACACGATCTGCTTAGCAATATGACAAAAGATGCTTTT	2361
QY	540	TCTCAAAACCCCGCTCTTTTACCAGAAATACAAAGGCACTAAAGCAACCCCTTTGGAAA	599
DB	2362	TCTCAAAACCCCGCTCTTTTACCAGAAATACAAAGGCACTAAAGCAACCCCTTTGGAAA	2421
QY	600	AGGTTGCAAGAATAGAGCGCGCGAGTACATTCCTTTCTATCAACAAGATTCTCAT	659
DB	2422	AGGTTGCAAGAATAGAGCGCGCGAGTACATTCCTTTCTATCAACAAGATTCTCAT	2481
QY	660	AACAAGACTTTTAACTTGAAGTCTGCTAGTACAGTTCAATTTGCTTCAGTCATTCGCAAG	719
DB	2482	AACAAGACTTTTAACTTGAAGTCTGCTAGTACAGTTCAATTTGCTTCAGTCATTCGCAAG	2541
QY	720	GAAGAGCTCAGCCATGTGTGCAA	742
DB	2542	GAAGAGCTCAGCCATGTGTGCAAGTATATACATATAGCCTCAACCAATTAATTA	2601
QY	743	-----	742
DB	2602	GTTTTATTCTTTCAAAATTTTATATTTTATTTCTATGCTCTTTTGGTTTCTTATTCT	2661
QY	743	-----	742
DB	2662	CATTATTAATCTATCAGTATCTACAAATTAAGGAAAGAAATAAGCTAACAAAAACATA	2721
QY	743	-----ATGGTGGAAAGCTTTTGATATCAAGAAAGACGCACTTTGTT	783
DB	2722	ATTATGATGTTACATAGGTGGTGGAAAGCTTTTGATATCAAGAAAGACGCACTTTGTT	2781
QY	784	TAAAGATAGAAATTTGAAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGT	843
DB	2782	TAAAGATAGAAATTTGAAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGT	2841
QY	844	ATTCGGGCTAGAGTTTCTTCAAAAAGCTTTGCTGTTTAACTCTTATAGATGACA	903
DB	2842	ATTCGGGCTAGAGTTTCTTCAAAAAGCTTTGCTGTTTAACTCTTATAGATGACA	2901
QY	904	CTTATGATCGGTATGGTACTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAA	959
DB	2902	CTTATGATCGGTATGGTACTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAA	2961
QY	960	-----	959
DB	2962	ATTTGTTAAACATGCTATCTGTCACACCTTACACACATACATATATTAACATA	3021
QY	960	-----	959
DB	3022	GAACCTTCAATTTAAACCCCATGTATAAATCAAACTAAATACGCAAGTCTCAATTA	3081
QY	960	-----AGGTGGTCAATTTACATGCTTAGACACACTTCCAGAAATACATGAAACCGATATACAA	1015
DB	3082	CGACAGGTGGTCAATTTACATGCTTAGACACACTTCCAGAAATACATGAAACCGATATACAA	3141
QY	1016	ATTATTCATGGATACATACAGAAATGGAAGAAATTTCTTCAAAAGGAGGGAAGACAGA	1075
DB	3142	ATTATTCATGGATACATACAGAAATGGAAGAAATTTCTTCAAAAGGAGGGAAGACAGA	3201

QY	1076	TCTATTTAACTGGCGCAAGAAATTT	1100
DB	3202	TCTATTTAACTGGCGCAAGAAATTTGTAAGTATCAACACCAATCTTTGATATTTTAAAA	3261
QY	1101	-----	1100
DB	3262	AAAAAATCCTGATGAAGATAATATATGTAAGACACGCTGACTGTATGCTATGTCATGCAC	3321
QY	1101	--GTGAAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGAGGACACA	1158
DB	3322	AGGTGAAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGAGGACACA	3381
QY	1159	TACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTCTAACCTGCTTA	1218
DB	3382	TACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTCTAACCTGCTTA	3441
QY	1219	CAACAACTTGTATCTTGGCATGATATATTTCAAAAAGAGTCTGTGGAATGGGCTG	1278
DB	3442	CAACAACTTGTATCTTGGCATGATATATTTCAAAAAGAGTCTGTGGAATGGGCTG	3501
QY	1279	TCTCTGACCTCTCTTTTGTAGATACATGATATCTTGTGTCGAGCCCTAAATGATCTCA	1338
DB	3502	TCTCTGACCTCTCTTTTGTAGATACATGATATCTTGTGTCGAGCCCTAAATGATCTCA	3561
QY	1339	TGACCCCAACA	1347
DB	3562	TGACCCCAACAGGTACACCCCTTTTCTAGTAGCATATCTATTAATTTCTGTTCT	3621
QY	1348	-----	1347
DB	3622	TAAACTCTAGTTTATGCCCCGAGCCTAAATAACCAATTTATTTTGTAAATATAATTTAT	3681
QY	1348	-AGGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTATATGAAGAAATAT	1406
DB	3682	CAGGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTATATGAAGAAATAT	3741
QY	1407	AATGTCAATGAGAGATATGCCAAACCTTGATTTTACAAGGAAGTAGAAGATGTGTGAAA	1466
DB	3742	AATGTCAATGAGAGATATGCCAAACCTTGATTTTACAAGGAAGTAGAAGATGTGTGAAA	3801
QY	1467	GATATAAACCGAGAGTACCTCACAACTTAAACCAATTTCCAGAGCCGTTATTGATGCTGTG	1526
DB	3802	GATATAAACCGAGAGTACCTCACAACTTAAACCAATTTCCAGAGCCGTTATTGATGCTGTG	3861
QY	1527	ATCTATTTGGCCAGTTTCTTGAAGTTCAATATGACAGAAAGGATTAATTCACAGTATG	1586
DB	3862	ATCTATTTGGCCAGTTTCTTGAAGTTCAATATGACAGAAAGGATTAATTCACAGTATG	3921
QY	1587	GGAGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTATCTATGAGTATATGAGA	1646
DB	3922	GGAGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTATCTATGAGTATATGACTA	3981
QY	1647	TC 1648	
DB	3982	CC 3983	

RESULT 12

AAN249561
LOCUS AAN249561 1893 bp mRNA linear PIN 23-SEP-1999
DEFINITION Artemisia annua mRNA for putative sesquiterpene cyclase.
ACCESSION AJ249561
VERSION AJ249561.1 GI:5921216
KEYWORDS sesquiterpene cyclase.
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroidae;
Anthemideae; Artemisia.
REFERENCE 1
AUTHORS Van den Beekhout, E., Van Geldre, E., De Pauw, I. and Inze, D.

TITLE Cloning sesquiterpene cyclase from *Artemisia annua*
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1893)
AUTHORS Van Geldre, E.G.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-1999) Van Geldre E.G., Laboratory for
Pharmaceutical Biotechnology, University of Ghent, Harelbekestraat
72, 9000 Gent, BELGIUM
CDS.

FEATURES
source Location/Qualifiers
1..1893
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
34..1877
/codon_start=1
/evidence=experimental
/product="putative sesquiterpene cyclase"
/protein_id="CAB56499.1"
/db_xref="GI:5921217"
/db_xref="GOA:Q9ST45"
/db_xref="SPTREMBL:Q9ST45"

CDS
translation="MSLIVEDVIRNPANFPSSIWGDFLAYDQDSQEGVEVQIKLKE
EVKSELLTALNSPTQHTSLLFIDAIERLGIAYFEESINQVFOHMYTAYGDKWTGGN
TSLFRLMRQHFVSSDIFSTYKDKGRFKESLEKDVHGLLELYEAAYMFVPGGIL
DDALVFPTCLDERAKNPSLNSAISLSOIREALTOPLHKLRLPRLEALRYIPFYQOAS
HSEILLKALGFNQLSHKELSLISKNWKSFEVANNLPVARNRPVCEYEWALAVY
PEPOYESRVLFSRFSIQTFDDYDAYGTTEELEOFTEALQKRSITCLDGLPSMK
LIFQWLKIFEEIEILSKDGQKHVNIKTELKEAVOSYMEARWAKEEYIPTIEEH
TKVSYISGYKALVAGFCMGDVIADDSFEWFTNPLVNACCLLCRTMDLGSKKG
EQDKHVASITCEYMKQFDASQOQAYESLNKKVEDAWKEINREFMITCKDVNIHVAMR
VLNFSRSVDVLYKNKDHETHVGEVINHKLSLFDVAIIT"

ORIGIN

Query Match 43.8%; Score 721.8; DB 8; Length 1893;
Best Local Similarity 65.3%; Pred. No. 1.8e-155;
Matches 1075; Conservative 0; Mismatches 567; Indels 3; Gaps 1;

QY 2 CATGCCACTTACAGAGAAAACCTTATTCGCCCATTCGCCAATTCCTCCCAAGCATTTG 61
DB 33 CATGCTCTTATAGTAGAGATGTCTATCGCCCAATGCAAAATTTTCCTTCTGAAATTTG 92
QY 62 GGGAGATCATGTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTAA 121
DB 93 GGGAGATCATGTTCTGCTATGACAGATGAGCAAGGGAGTGTGAACAAGTATCA 152
QY 122 TGATTTAAAAAGAAAGTGGCGCACTACTAAAGAAAGCTTTGGATATTCCTATGAACA 181
DB 153 AGATTTGAAGAAAGAAAGTGAAGAGCAACTATTGACAGCTCTAAATTCCTCGACTCAGCA 212
QY 182 TGCCAAATTTGTAAGCTGATGATGAATTAACCGCTTGGATACCGTATCACTTTGA 241
DB 213 TACGGAGTTGTTGAAATTTATTGACGCAATCGAACCGCTTTGGTATTCGGTATTTTGA 272
QY 242 ACGGGAGATTGATCATGCTGCTTCAATGTATTTATGAACAATATGGTGATTAACCTGGAATGG 301
DB 273 AGAGGAGATCAACCAAGTATTCACACATGTATCTACTGATATGGTATGATGAGCAGG 332
QY 302 TGACCGCTCTTCTCTTATGGTTCGGTCTTATCGAAAGCAAGGATATTATGTTACATGTGA 361
DB 333 TGGTAACTCTCTCTTGGTTTGGCTCATGCGCAACATGGAATTTTCGTTTCAAGCGA 392
QY 362 TGTCTTCAATACTATAAGACAAATAATGAGCGCTTCAAGCAATCGTTAGCTATGATGT 421
DB 393 TATTTTGTAGTACCTACAAAGCAAGAGGGACGTTTAAAGAGTCTTTAGAAAGGAGCGT 452
QY 422 TGAAGTTTCTGCTGAGTTGTACGAGCAACTTCTATGAGGCTACTCGGGAGATATTATTT 481
DB 453 TCACGGGTTTACTTGTAGTTGTATGAGCGGCTATATGTTTGTGCTGGGAGGAGTACT 512
QY 482 AGAAGATGCTCTTGGTTTACACGATCTCGTCTTAGCATATTATGACAAAGATGCTTTTC 541
DB 513 AGATGATGCTCTGTTTTCAGAAAGACTTGTCTTGTGAATAGCAAAATCTCTAGTCT 572

QY 542 TACAACCCCGCTCTTTTATCCGAAATACAAACGGGCACTAAAGCAACCCCTTTGAAAAAG 601
DB 573 AGTAACTCTGCTATTTTCTTCCCAATACGTGAAGCTCTAAACGCAACCTTTTACATAAAG 632
QY 602 GTTGCCAAAGATAGAGCGCGCGAGTACATTCCTTTCTATCAACAAAGATTTCTCATAA 661
DB 633 ATTACCGAGACTAGAGCGCTTGGCTACATACCTTTCTACCAACAAAGCTTCTCAG 692
QY 662 CAAGACTTTACTTTAAACTTGTCTAAAGTTAGAGTTCAATTTGTCTCAGTCAATTCGCAAGGA 721
DB 693 TGAACATTTGTTAAACTGSCCAATTTAGGGTTCAACCAACTTCAATCCTTCAACAGAA 752
QY 722 AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCATATCAAGAAAGACGACCTTG 781
DB 753 AGAGCTTAGCATAAATTTCCAAAGTGGTGGAAAGTTTCGATGTTGCAAAACAATCTACCGTA 812
QY 782 TTTAAGAGATAGAAATGTTTGAATGCTACTTTTGGGACTAGTTTCAGGCTATGAGCCACA 841
DB 813 TGAAGAAATAGACCCGCTTGAATGCTACTTTTGGGCAATTGACAGTGTACTTTCGAGCCTCA 872
QY 842 GTATTCGCCGGCTAGAGTTTCTTCAAAAAGCTTGTCTGTATATACTTATATAGATGA 901
DB 873 ATATTCGAATCAAGAGTCTTCTTATCAAGATTTTTCATACAAACAATTTCTTGTATGA 932
QY 902 CACTTATGATGCGGTATGCTACTTATGAAGAACTTAAAGATCTTTTACTGAAGCTGTTGAAAG 961
DB 933 CACTTATGATGCGGTACTTACGAAAGAACTTGAACAATTTTACTGAAGCAATTTCAAAAG 992
QY 962 GTGGTCAATTACATGCTTAGACACACTTCCAGATACATGAACCGCATATACAAATTTAT 1021
DB 993 GTGGTCAATAACATGCTTGGACGGTCTTCTGAAAGCATGAAGCTTATATTCCAAATGCT 1052
QY 1022 CATGGATACATACAGAAATGGAAGAAATTTCTTCAAAAGGAGGGAAGCAAGATCTATT 1081
DB 1053 AGTGAAGATTTTGAAGAAATAGAGGAATATTGTGAGAGGTGGAACACATCATGT 1112
QY 1082 TAACTGCGGCAAGAAATTTGTGAAGAGTTTGTGAAGAACTGTATGAGTGGTTGAAGCAAAATG 1141
DB 1113 CACTACATCAAGAAACACCTGAAAGAGGCTGTCCAAAGCTACATGACAGAAAGCAAGATG 1172
QY 1142 GGCATAATGAGGACACATACCAACCACTGAAGAGCATGNTCCAGTTGTATCATTTACTGCG 1201
DB 1173 GGCATAAGAGAGATATATACCAACGATTGAGAGACACAAAGGTGTATATATTAGCAT 1232
QY 1202 CGGTCTAACTGCTTACCAACAACTTGTATCTTGGCATGAGTGATATATTCAAAAGA 1261
DB 1233 CGGCTATAAATTTGCGCTAGTAGAGGTTTGTGTCATGGTGTATTTGCAGATGA 1292
QY 1262 GTCTGTGGAATGGGCTGTCTGCACTCTCTTTTAGATACACTCAGGTATCTTGGTTCG 1321
DB 1293 CTCATTTGAATGGGTTTTCACCAATCTCTCTTGTCAATGCTTGTGTCTACTTTGTAG 1352
QY 1322 ACGCTAAATGATCTCATGACCCCAAGCGCGAGCAAGAAAGAAACATAGTTTCATCGAG 1381
DB 1353 AACCATGGATGATCTTGGTTCCTATAGGGCGAGCAAGATAGAAAGCATGTTGCTCTCAC 1412
QY 1382 CTTTGAAGATTTATGAAGGAATAAATGTCATGAGGAGTATGCCCAACCTTTGATTTA 1441
DB 1413 TATTGAATGCTACATGAAGCAATTTGATCTAGCGAGCAACAGCTACGAATCACTCAA 1472
QY 1442 CAAGGAAGTAGAAGATGTTGGAAGAGATATAAACCGAGAGTAC---CTCAACTAAAAA 1498
DB 1473 CAAGGAAGTTGAAGATGATGGAAGAAATAAATAGAGAGTTTATGATCAGTGTAGGA 1532
QY 1499 CATTCGAAGCGGTTATTTGATGCTGCTGATCTATTTTGTGCGCAGTTTCTTGAAGTTCAATA 1558
DB 1533 TGTAAATATTCATGAGCTATGAGGCTCTTAAATTTTCACTGTTGGTGGATGATGCTCTTTA 1592
QY 1559 TGCAGGAAGGATAACTTCCACAGTATGGAGAGCAATAAACAATCTCTATAAAGTCTCT 1618
DB 1593 TAAGAAATAGGATCACTTTTACATGTTGGAGTAGAGGTGATAAATCATATCAATCTCT 1652

```
QY 1619 ACTCGTTTATCCTATGAGTATATGA 1643
Db 1653 TTTCGTTGATGCCATCATATCATGA 1677

RESULT 13
AF304444 AF304444 1886 bp mRNA linear PLN 23-FEB-2001
LOCUS Artemisia annua sesquiterpene cyclase mRNA, complete cds.
DEFINITION Artemisia annua sesquiterpene cyclase mRNA, complete cds.
ACCESSION AF304444
VERSION AF304444.2 GI:131113759
KEYWORDS
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
1 (bases 1 to 1886)
Liu Y., Ye H.C. and Li, G.F.
Cloning, characterization and E.coli expression of a cDNA encoding
sesquiterpene cyclase from Artemisia annua
2 (bases 1 to 1886)
Liu Y., Ye H.C. and Li, G.F.
Direct Submission
Submitted (11-SEP-2000) Department of Cell and Gene Engineering,
Institute of Botany, Chinese Academy of Sciences, No. 20 Nanxin
Chun, Xiangshan, Beijing 100093, P.R. China
3 (bases 1 to 1886)
Liu Y., Ye H.C. and Li, G.F.
Direct Submission
Submitted (23-FEB-2001) Department of Cell and Gene Engineering,
Institute of Botany, Chinese Academy of Sciences, No. 20 Nanxin
Chun, Xiangshan, Beijing 100093, P.R. China
Sequence update by submitter
On Feb 23, 2001 this sequence version replaced gi:10945664.
FEATURES
location/Qualifiers
1..1886
/organism="Artemisia annua"
/mol_type="mRNA"
/strain="001"
/db_xref="taxon:35608"
/country="China: Sichuan"
42..1763
/codon_start=1
/product="sesquiterpene cyclase"
/protein_id="AAG24640.2"
/db_xref="GI:13113760"
/translation="MSTFLVCNAPSFPSPVCLPSSCDNSNMQEIYRNTVKPPSIWG
DPLITHEPKSLTIKQHVVEEIRKLLNTAFSEPKOHTKLHLIDSVQRLGVSY
HFEQTEELALQHYATHGDQWIGKNLKSTQWFRILRQQGFNVSSGIFENHMDKQK
FKEHLSDDVQGMGLYEAYMSVEGEKELDDALEFTKHLGNIAKOPSONASLRKIE
QALNRPLRMRPLREALHYIPKYQOASHDETLNLAKLDYNNLQSIHKREI SEICKW
WDLDPSNKLPHVRDLRLVETFWLGVYFEPQVSRMPLTKTSMMLIVLDDYDNYG
TYEELKIFTEAVORSLCLDSLPEYMKLIYRELINHYQWESSEVEKESYQLFYAK
EMVKNRMLLVEAKKEGVFTLEHMSVSCVTYAYAVIANSYVRENMYTESF
KWAKFPPLVLNMDCLLIRLMDIDTAGHEEQRDHVSSIECYRNETGASEEDAVKFLS
KQVEDAKVINKECLEPTEIPMSLTIPPLSLARVSDILYKTNNGYNHAGEVISYIKS
LLVHPLV"

CDS
Query Match 35.7%; Score 588; DB 8; Length 1886;
Best Local Similarity 61.9%; Pred. No. 1.le-124;
Matches 1011; Conservative 0; Mismatches 600; Indels 23; Gaps 4;

QY 28 TTGCGCCCATTCGCAACTTTCCTCCAGCATTTGGGGAGATCAGTTTCTCATCTATCA-- 85
Db 136 TCGGCAACTGTATAATTTCTTCTAGCATATGGGGAGATCAGTTTCTTACATACCATG 195
QY 86 -----AAAGCAAGTAGCAAGGGGTGGAAACAGATAGTGAATGATTTAAAAAAGAAAGTG 140
Db 196 AACCGAAAAAAGTTTAAACCATTTGAGAAACAAACATGTTGAAGAGGTTGAAGAAATAAG-G 254
```

```
QY 1209 AACCTGCTTACAACTGTTATCTTGGCATGAGTATATATACAAAAAGAGTCTGTC 1268
Db 1332 GCTGTCAATGTCGAACTCTTATGTTGGACGGGAAATATGTTTACAGAGGATCCCTTT 1391
QY 1269 GAATGGGCTGCTCTGACCTCTCTTTTATAGATCTCAGGTATATCTTGGTCAGCCCTA 1328
Db 1392 AAGTGGGTTGCCAAATTTCTCTCTTGTAAATGCTACATGTTTATGTTTAAAGGCTTATG 1451
QY 1329 AATGATCTCATGACCCACAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCCTTGA 1388
Db 1452 GATGATATCGGGCCACGAGGAGAAACAGAAAGAGACCATGTTGTTTCAAGTATTGAA 1511
QY 1389 AGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTCAATTTACAGGAA 1448
Db 1512 TGCTACAGAAATGAACTGGAGCTCGAGAGGAGTGGCTTAAGTTTCTTTCAAAACA 1571
QY 1449 GTAGAGATGTTGGAAGATATAAACCAGAGTACCTCAACAATTAACAACTTCAAGG 1508
Db 1572 GTTGAGGATGATGGAAGATTATAAATGAAGAGTGTCTAAGGCTTACAGAAATTTCCCATG 1631
QY 1509 CGTTATATGAGGCTGATCTATTTGTCAGGTTTCTTGAAGTTCAATATCGAGGAAG 1568
Db 1632 TCTCTAACTATACCTCCACTCAGCCTTCTCTGCTGAGTGATATCCTTTATAAAACGAAT 1691
QY 1569 GATAAATTCACAGTATGGGAGAGCAATACAAAACATCTCAAAAGTCTCTACTCGTTTAT 1628
Db 1692 AATGTTACAAATCAGCAGGGGAGAGTATGATCAGTTACATTAAGTTCGCTCTGTTTCAAT 1751
QY 1629 CCTATGATGATG 1642
Db 1752 CCTTTAGTTTAATG 1765

RESULT 14
AAN271792 1843 bp mRNA linear PLN 04-OCT-2000
DEFINITION Artemisia annua mRNA for putative sesquiterpene cyclase (cASC125 gene).
ACCESSION AJ271792
VERSION AJ271792.1 GI:10696987
KEYWORDS cASC125 gene; sesquiterpene cyclase.
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
REFERENCE 1
AUTHORS Van Geldre,E., De Pauw,I., Inze,D., Van Montagu,M. and Van den
Beckhout,E.
TITLE Cloning and molecular analysis of two new sesquiterpene cyclases
from Artemisia annua L
JOURNAL Plant Sci. 158 (1-2), 163-171 (2000)
PUBMED 10996256
REFERENCE 2 (bases 1 to 1843)
AUTHORS Van Geldre,E.G.L.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Van Geldre E.G.L., Laboratory for
Pharmaceutical Biotechnology, University of Ghent, Harelbekestraat
72, 9000 Gent, BELGIUM
FEATURES
source
location/Qualifiers
1..1843
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/tissue_type="leaves"
1..1843
/gene="cASC125"
61..1794
/gene="cASC125"
/codon_start=1
/evidence=experimental
/product="putative sesquiterpene cyclase"
```

```
protein_id="CAC12731.1"
/db_xref="GI:10696988"
/db_xref="GOA:Q9FY7"
/db_xref="SPTREMBL:Q9FY7"
/translation="MSTFLPISSVSFSSTSLVVDVDDKVKSTKPDVIRHTMTNINSLWGD
QFLTDEPDLVMKQLVBEKEVKELITKSNFPMQHVKLIELIDAVQVGLGIAY
HFEETEEALQIHVITYGQWVDKENLQISLWFLRLQQQFNYSVGFQKDFMDEKGIH
FRESLNDAGIILALVEAAMRVEDETILDNALEFTKVHLDIIAKDPSCDSLSLQIKH
QALKPLRLARIEALHVMPIYQOETSHNEDLLKAKLDFSVLQSMHKKLSHICKW
WKDLQNKLPVVRDRVVEGYFWILSIYEPQHARTFMELMKTOMLVLLDDTPNYG
TYEELIFTQAVVERWSICLMDPEYMKLIYQELVNLHVEEMESLGKGRKINISLQC
GWQKELGSOITLIVETMAKRGVHAQPLEEYMSVMTGTGLMIARSYVGKGDIVTE
DTFKWSSVYPIIKASCVIIRLMDDIVSHKEQERGHVASSIECYSEKSGASEBBEAC
YTSRKVEDAWKVINRESLRPTAVFPFLMLPAINLARMCEVLVSNDGFTHAEGDMKSY
MKSFFVHPVWV"
ORIGIN
Query Match 33.9%; Score 559; DB 8; Length 1843;
Best Local Similarity 61.6%; Pred. No. 5.2e-118;
Matches 1025; Conservative 0; Mismatches 605; Indels 33; Gaps 7;
QY 12 ACAGAAAGAAACCTATTGCCCCCATGCCAACTTTCTCCCAAGCATTTGGGGAGATCAG 71
Db 136 ACAGAAACCGATGTTTATCGCCATACCATGAATTTCAATGCTAGCATATGGGGAGATCAA 195
QY 72 TTTCTCATCTATCAAAAGCAAGTAGCAAGGGTGG---AACAGATAGTGAATGATTTA 128
Db 196 TTTCTTACATATGATGAGCCAGAGATTGGTTATGAAGAAACAACCTGTTGTAAGAGCTG 255
QY 129 AAAAAAGAAAGTCGGCAACTTACTATAAAG-----AAGCTTTGGATATTCCTATGAAACAT 182
Db 256 AAGAGGAAAGTCAAAAAGAGCTTAATAACGATCAAGAGTTCAAAATGAACCGATGAGCAT 315
QY 183 GCCAATTTGTTGAAGCTGATTGATGAATAATTCACGCCCTTGGAAATACCGTATCAGTTTGA 242
Db 316 GTCAAGTTGATAGAACTTATTGATGCAAGTCCACGGCTTGGCATAGGCTATCATTTTGA 375
QY 243 CGGGAGATTGATCATGATTCGCAATGATTTATGAAACATATGTTGATAAATCTG----- 296
Db 376 GAGGAGATTGAAGAGCCTTGCAACATATCCATGTTACATATGCGGAGCAGTGGTTCAT 435
QY 297 -----AATGGTGACCGCTCTTCTTATGTTTCCGTCTTATGCGAAAGCAAGGATATTAT 350
Db 436 AAGAAACCTACAAAGCATTTCTTTGTTGCTCGACTCTCTACGACAAAGGCTTCAAC 495
QY 351 GTTACATGATGTTTTCATTAATCTATAAGACAAATAATGAGGCGTTCAAGCAATCGTTA 410
Db 496 GTCTCATCAGGAGTATTCAAAGACTTTATGACGAGAAAGGAAAAATTTAAGGAATCCTTA 555
QY 411 GCTAATGATGTTGAAGGTTTGGCTTGAAGTTGAGTTGACGAAGCACTCTATGAGGATCCTGG 470
Db 556 TGTAATGATGCTCAGGGAATCTTGTCTGTATGAAGCGGCATTTATGAGGTTGAAGAT 615
QY 471 GAGATTATATTAGAAGATGCTCTTTGTTTACAGATCTCGTCTTAGCATTATGACAAA 530
Db 616 GAAACAATACTAGATAATGCGCTCGAGTTCAATAAGTTCACTACCTTGACATATAGCCAG 675
QY 531 GATGCTTTTCTACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAGCAACCC 590
Db 676 GATCCTTCTTGT---GACTCTTCCCTAAGAACCCAAATACACCAAGCGCTTAAGCAGCCA 732
QY 591 CTTTGAAGAGGTTGCCAAGAAATAGAGCGCGGCGAGTACATCTCTTCTATCAACAAACA 650
Db 733 CTTAGGAGAGGCTGGCAAGGATGAGCGGTGTCATTATCATGCTTATCTACCAACAGAA 792
QY 651 GATTTCTATAACAAGACTTTTACTTAACTTGAAGTTAGAGTTCAATTTGCTTCAGTCA 710
Db 793 ACATCCCAATATGAGGACTTACTGAAGCTTGCAAAATTAGATTTTAGCGTTCAGTCG 852
QY 711 TTGCACAAAGAGAGCTCAGCCATGTTGCAAAATGTTGGTGAAGAGCTTTCGATATCAGAG 770
Db 853 ATGCACAAAGAGAACTTAGCCACATTTGCAAAATGTTGGTGAAGAGATTTGAGCCTCCAAAT 912
```


771 AACGACCTGTTAAGAGATAGAAATGTTGAATGCTACTCTTTGGGACTAGGTTGAGC 830
Db AAGCTTCCTTTATGTTTCGAGACAGGGTGGTGAAGGCTATTTTGGATACTACCTCTAT 972
Qy TATGAGCACAGTATTCGCGGCTAGAGTTTCTTTCACAAAGCTGTTGCTGTATACT 890
Db TACGAGCTCAACATGCTCGAACAAAGATGTTCTTAATGAAACATGATGTTGTTAGTC 1032
Qy CTTATAGATGACACTTATGATGCTGATGCTATGATGCTTATGAAAGACTTTAAAGTCTTTACTGAA 950
Db GTTTTAGATGATACATTTTGATAATTTATGCTATGCTTATGAAAGACTCGAAATCTTTACACA 1092
Qy CTGTTGAAAGTGTGCTAATACATGCTTAGACACATCTCCAGATACATGAAACCGATA 1010
Db GCTGTTGAAAGATGCTCAATAAGCTGCTCGGATATGCTTCCAGAAATACATGAACTAAATA 1152
Qy TACAAATATTATCATGATACATACACAGAAATGGAAGATTTCTT-GCAAAGGAGGAG 1069
Db TATCAGAGCTTGTAATCTTCACGTGGAATGGAAGATCACTTGGAAAGAGGAGAA 1212
Qy AACAGATCTATTTAACTCGGCAAGAAATTTGTGAAAGA-----GTTTGTTAGAAACCT 1123
Db AACATATCAATTTCACTATGTCAAGGAGATGCGAAAGAGTGGGTTTCGCAATTAATCT 1272
Qy GATGTTGAAAGAAATGGGCAATGAGGACACATACC--AACCACTGAAAGCATGAT 1181
Db TTGTAGAAACCAAGATGGCTAAAAGAGGGGTACATGCGCCACCTCTTTGAGGAGTACATG 1332
Qy CCAGTTGTAATCATTTACTTACCTGCTTACCAACCTGTTTATCTTGGCATG 1241
Db TCTGTCAATGGTCACTGATGCTATGCTTATGATGATGAGGATCTTATGTCGCAGG 1392
Qy AGTGATATTTACAAAAGAGTCTGCGAATGGGCTCTCTGCACTCTCTTTTATG 1301
Db GGTGATATCGTCACCGAGATACCTTTTAAATGGTGTCTCTGATCTCTCTATATAAAA 1452
Qy TACTCAGTATATCTTGTGCGAGCTCAATGATCTCATGCCACCAAGCCGAGCAAG 1361
Db GCTTCATGTTGATTTAAGCTTTATGATGATTTGTCAGCCACCAAGGAGCAAGAG 1512
Qy AGAAACATAGTTTCATCGAGCTTGAAGTATATGAAAGAAATATAATGTCATGAGGAG 1421
Db AGAGCCATGTTCTCAAGCATGCAATGCTATTTCAAGAAAGTGGTGCATCAGAGGAG 1572
Qy TATGCCAAACCTTGATTTACAGGAAGTAGAAGTGTGTGGAAGATATAAACCGAGAG 1481
Db GAAGGCTGTAATATATCTCAAGAAAGTGTGAAGTGCATGGAAGCTTATAAACCGAGAG 1632
Qy TACCTCACAACCTAAAACATTCAGAGCCGTTATGATGCTGTGATCTATTTGTGCCAG 1541
Db TCGCTCAGGCTACAGCTGTCCCATTTCTCTACTATGCTGCAATCAACCTTGACCG 1692
Qy TTTCTTGAAGTTCAATATGCGAGGAAGATAACTTCACACGCTATGGGAGACGAATACAA 1601
Db ATGTGGAAGCTTATATAGCTCAAGATGCTTCACATGCTTCATGCTGAGGAGACATGAA 1752
Qy CATCTCATAAGCTCTACTCTGTTTATCTTATGATGATATGAG 1644
Db AGTTACATGAATCTGTTCTTCTGTTTACCCCTATGTTGTTCTAAG 1795

RESULT 15
BD227439
LOCUS BD227439 538 bp DNA linear PAT 17-JUL-2003
DEFINITION Transgenic amorpho-4,11-diene synthesis.
ACCESSION BD227439
VERSION BD227439.1 GI:33037209
KEYWORDS JP 2002523101-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1. (bases 1 to 538)
AUTHORS Wallaart,T.E. and Bouwmeester,H.J.

Transgenic amorpho-4,11-diene synthesis
Patent: JP 2002523101-A 1 30-JUL-2002;
GENOLIPP BIOTECHNOLOGY BV
OS Artemisia annua L.
PN JP 2002523101-A/1
PD 30-JUL-2002
PF 27-AUG-1999 JP 2000567711
PR 27-AUG-1998 EP 98202854.0
PI THORVALD BELCO WALLAART,HENDRIK JAN BOUWMEESTER PC
C12N15/09,C12N15/09,A01H5/00,C12N1/19,C12N1/21,C12N5/10,C12N9/ PC
88,C12P5/00,
PC C12P17/18,C12N9/04,C12N1/19,C12R1/84),(C12N1/19,C12R1:645),
PC (C12N9/04,C12R1:91),(C12N9/04,C12R1:865),(C12N9/04,C12R1:84),
PC (C12N9/04,C12R1:19),(C12N9/04,C12R1:645),(C12N9/88,C12R1:91),
PC (C12N9/88,C12R1:19),(C12N9/88,C12R1:84),(C12N9/88,C12R1:865),
PC (C12N9/88,C12R1:645),(C12N15/09,C12R1:91),(C12P5/00,C12R1:19),
PC (C12P5/00,C12R1:84),(C12P5/00,C12R1:91),(C12P5/00,C12R1:865),
PC (C12P5/00,C12R1:645),(C12P17/18,C12R1:645),(C12P17/18,C12R1:19) PC
PC (C12P17/18,C12R1:91),(C12P17/18,C12R1:84),(C12P17/18,C12R1:865) PC
C12N15/00,
PC C12N15/00,C12N5/00,(C12N15/00,C12R1:91)
CC PCR product amplified on cDNA as template with primers CC
CC starting and ending
at nucleic acid positions (1..23 and 517..538) FH Key
Location/Qualifiers
FT CDS (1)..(537).
source 1..538
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 31.08; Score 510.6; DB 6; Length 538;
Best Local Similarity 97.4%; Pred. No. 7.9e-107;
Matches 519; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 381 GACAAAATCGAGCGTTCAGACCAATCGTTAGCTAATGATGTTGAAGGTTGCTTGAGTTG 440
Db 1 GATGAGATCGGAAATTTAAGGAATCGTTAGCTAATGATGTTGAAGGTTGCTTGAGTTG 60
Qy 441 TAGAAGCAACTCTCTATGAGGTACCTGGGGAGATTATATTAGAAGATGCTCTTGTTT 500
Db 61 TAGAAGCAACTCTCTATGAGGTACCTGGGGAGATTATATTAGAAGATGCTCTTGTTT 120
Qy 501 ACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCGCTCTTTT 560
Db 121 ACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCGCTCTTTT 180
Qy 561 ACCGAATACACGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGATAGAGCG 620
Db 181 ACCGAATACACGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGATAGAGCG 240
Qy 621 GCGAGTACATCTCTTCTATCAACAACAGATTTCTATACAGACCTTACTTAACTT 580
Db 241 GCGAGTACATCTCTTCTATCAACAACAGATTTCTATACAGACCTTACTTAACTT 300
Qy 681 GCTAAGTTAGAGTTCAATTTGCTTCACTTGCACAGGAAGAGCTCAGCCATGTGTC 740
Db 301 GCTAAGTTAGAGTTCAATTTGCTTCACTTGCACAGGAAGAGCTCAGCCATGTGTC 360
Qy 741 AAATGTTGAAAAGCTTTTCGATATCAAGAAGAACCCCTTTTAAAGAGATAGAATTGT 800
Db 361 AAATGTTGAAAAGCTTTTCGATATCAAGAAGAACCCCTTTTAAAGAGATAGAATTGT 420
Qy 801 GAATGCTACTTTTGGGCACTAGGTTGCTATGAGCCACAGTATTCGGGCTAGAGTT 860
Db 421 GAATGCTACTTTTGGGCACTAGGTTGCTATGAGCCACAGTATTCGGGCTAGAGTT 480
Qy 861 TTCTTCACAAAGCTGTTGCTGTTTAACTCTTATAGATGACACTTATGATGC 913

Db 481 TTCTTCACAAAGCTGTGCTGTATATAACTCTTATAGACGACACCTTCGACGC 533

Search completed: June 7, 2004, 11:23:28
Job time : 6517 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 08:48:09 ; Search time 735 Seconds
(without alignments)
9530.991 Million cell updates/sec

Title: US-09-763-822A-13

Perfect score: 1649

Sequence: 1 ccattgcattacagaagaa.....ctatgagtatataggatgcc 1649

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001as.*

5: Geneseq2001bs.*

6: Geneseq2002s.*

7: Geneseq2003as.*

8: Geneseq2003bs.*

9: Geneseq2003cs.*

10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	1649	100.0	1649	3	Aaz92754 Artemisia
2	1640.6	99.5	2112	3	Aaz92760 Artemisia
3	1637.4	99.3	2076	3	Aaz92763 Artemisia
4	1610.6	31.0	538	3	Aaz92757 Artemisia
5	396.2	24.0	1973	9	Ade82707 Terpenoid
6	369.8	22.4	2605	9	Ade82704 Terpenoid
7	364	22.1	1879	2	Aax89865 L. escul
8	363.6	22.0	1672	9	Ade82701 Terpenoid
9	357.6	21.7	2024	2	Aax89866 L. escul
10	357.6	21.7	2024	3	Aax89835 Tomato ge
11	330.4	20.0	1944	3	Aax89828 Potato ve
12	316.6	19.2	1644	3	Aax89813 Tobacco 5
13	315.4	19.1	1644	3	Aax89815 Tobacco 5
14	315	19.1	1644	3	Aax89810 Tobacco 5
15	311.8	18.9	1644	3	Aax89812 Tobacco 5
16	311.8	18.9	1644	3	Aax89811 Tobacco 5
17	311.8	18.9	1644	3	Aax89814 Tobacco 5
18	289.2	17.5	1755	3	Aax97448 Chicory g
19	257.8	15.6	1953	3	Aax97449 Chicory g
20	256	15.5	1994	3	Aax38930 Castor be
21	231	14.0	1699	3	Aac47081 Arabidops
22	225.6	13.7	1650	2	Aax35571 Computer
23	225.6	13.7	1650	2	Aax35572 Computer

24	225.6	13.7	1650	2	AAX35575	Aax35575 Computer
25	225.6	13.7	1650	2	AAX35574	Aax35574 Computer
26	225.6	13.7	1650	2	AAX35576	Aax35576 Computer
27	225.6	13.7	1650	2	AAX35578	Aax35578 Computer
28	225.6	13.7	1959	3	AAX38926	Aax38926 Mint E-b-
29	225.6	13.7	1960	2	AAX35570	DNA encod
30	224.8	13.6	1650	2	AAX35573	Aax35573 Computer
31	224	13.6	1650	2	AAX35577	Aax35577 Computer
32	223.2	13.5	1910	7	ADA71275	Ada71275 Rice gene
33	219.4	13.3	1976	3	AAC38691	Aac38691 Arabidops
34	217.2	13.2	1728	7	ADA68331	Ada68331 Arabidops
35	205.6	12.5	1524	7	ADA69380	Ada69380 Rice gene
36	204	12.4	1575	7	ADA70943	Ada70943 Rice gene
37	199.2	12.1	1631	9	ADE82680	Ade82680 Terpenoid
38	199.2	12.1	1874	9	ADE82682	Ade82682 Terpenoid
39	199.2	12.1	1874	9	ADE82678	Ade82678 Terpenoid
40	198.2	12.0	1665	9	ADE82745	Ade82745 Terpenoid
41	196	11.9	1665	9	ADE82747	Ade82747 Terpenoid
42	193.4	11.7	1894	9	ADE82684	Ade82684 Terpenoid
43	191.6	11.6	1560	7	ADA70344	Ada70344 Rice gene
44	183.6	11.1	1806	7	ADA69387	Ada69387 Rice gene
45	161.4	9.8	1878	3	AAC49451	Aac49451 Arabidops

ALIGNMENTS

RESULT 1

Aaz92754

ID Aaz92754 standard; cdna; 1649 BP.

XX AC Aaz92754;

XX DT 19-JUN-2000 (first entry)

XX DE Artemisia annua amorpho-4,11-diene synthase cdna.

XX KW Amorpho-4,11-diene synthase; amorphadiene synthase; FPP;

XX KW farnesyl pyrophosphate; cyclisation; quinghao; sweet wormwood;

XX KW annual wormwood; sweet annie; artemisin precursor; antimalarial;

XX KW sesquiterpene; flavouring; fragrance; ss.

XX OS Artemisia annua.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 3..1643

XX FT /*tag= a

XX FT /product= "Artemisia annua amorpho-4,11-diene synthase"

PN EP982404-A1.

XX PD 01-MAR-2000.

XX PP 27-AUG-1998; 98EP-00202854.

XX PR 27-AUG-1998; 98EP-00202854.

XX PA (WALL/) WALLAART T E.

XX PI Wallaart TED, Bouwmeester HJ;

XX DR WPI; 2000-258617/23.

XX DR P-PSDB; AAY81269.

XX PT New isolated DNA sequences and polypeptides comprising amorpho-4,11-diene

XX PT synthase activity, useful for production of amorphadiene and/or

XX PT artemisinin.

XX PS Claim 4; Fig 9A; 41pp; English.

XX CC This sequence represents cDNA encoding Artemisia annua amorpho-4,11-diene

XX CC synthase. This enzyme catalyses the cyclisation of farnesyl pyrophosphate

XX CC (FPP), a ubiquitous precursor of primary and secondary metabolites, into

authentic
acc
inverted

CC amorpha-4,11-diene. Amorphadiene is a precursor of artemisin, a
CC sesquiterpene lactone endoperoxide produced by A. annua which is a
CC promising antimalarial drug candidate. Amorphadiene synthase is a branch
CC point enzyme in the artemisin biosynthetic pathway, and the cyclisation
CC of FPP is therefore likely to be the rate-limiting step. Artemisin is
CC present in very low concentrations in A. annua, making it expensive for
CC use as an antimalarial drug, and attempts at organic synthesis have been
CC unsuccessful. Sesquiterpenes such as amorphadiene are also useful as
CC flavour and fragrance compounds in the food and perfume industries.
CC Terpenes also play a role in plant-insect interactions, such as the
CC attraction or repulsion of insects by plants. In addition,
CC dihydroarteannic acid, an intermediate in the metabolic route from
CC amorphadiene to artemisin in A. annua, can be used as an antioxidant.
CC DNA encoding amorphadiene synthase may be used to generate transgenic
CC plants able to synthesise amorphadiene synthase. Preferably, the plants
CC used for amorphadiene production are plants which naturally produce
CC sesquiterpenes, as these plants already have the basic sesquiterpene
CC synthetic pathways and storage compartments. A. annua may be transformed
CC with amorphadiene synthase expression constructs, enabling more
CC amorphadiene to be produced and hence increasing artemisin production. It
CC may then be economically feasible to extract artemisin from such plants
CC for use as an antimalarial
XX
SQ Sequence 1649 Bp; 537 A; 299 C; 350 G; 463 T; 0 U; 0 Other;

Query Match 100.0%; Score 1649; DB 3; Length 1649;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGCAGCTTACAGAGAAAACCTATTCCGCCCATTCGCCAATCTTCTCCCAAGCATTTT 60
DB 1 CCATGGCAGCTTACAGAGAAAACCTATTCCGCCCATTCGCCAATCTTCTCCCAAGCATTTT 60
QY 61 GGGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAAACAGATAGTGA 120
DB 61 GGGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAAACAGATAGTGA 120
QY 121 ATGATTTAAAAAAGAGTGGCGCACTACTAAAGAGAGCTTTGGATATTCCTATGAAC 180
DB 121 ATGATTTAAAAAAGAGTGGCGCACTACTAAAGAGAGCTTTGGATATTCCTATGAAC 180
QY 181 ATGCCAATTTGTTGAAGCTGATGATGAAATTCAGCGCTTGGATATCCGTATCACTTTG 240
DB 181 ATGCCAATTTGTTGAAGCTGATGATGAAATTCAGCGCTTGGATATCCGTATCACTTTG 240
QY 241 AACGGGAGATGTATCATGCAATTCGAATGATTTATGAACAATATGGTGATTAACGGAAATG 300
DB 241 AACGGGAGATGTATCATGCAATTCGAATGATTTATGAACAATATGGTGATTAACGGAAATG 300
QY 301 GTGACCGCTCTTCTTATGTTCCGTCTTATGCGAAGCAAGGATATTTATGTTACATGTG 360
DB 301 GTGACCGCTCTTCTTATGTTCCGTCTTATGCGAAGCAAGGATATTTATGTTACATGTG 360
QY 361 ATGTTTTCAATACTATAAGACAAAAATGAGCGTTCAAGCAATCGTTAGCTAATGATG 420
DB 361 ATGTTTTCAATACTATAAGACAAAAATGAGCGTTCAAGCAATCGTTAGCTAATGATG 420
QY 421 TTCAAGGTTTGTGAGTTGTACGAAACAACTTCTATGAGGGTACCTGGGAGATTATAT 480
DB 421 TTCAAGGTTTGTGAGTTGTACGAAACAACTTCTATGAGGGTACCTGGGAGATTATAT 480
QY 481 TAGAAGATGCTCTGTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTT 540
DB 481 TAGAAGATGCTCTGTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTT 540
QY 541 CTACAAACCCCGCTCTTTTACCGAAATACAAAGGCGCACTAAGCAACCCCTTTGGAAA 600
DB 541 CTACAAACCCCGCTCTTTTACCGAAATACAAAGGCGCACTAAGCAACCCCTTTGGAAA 600
QY 601 GGTGGCAAGAAATAGAGCGCGGCGAGTACATTCCTTTCTATCAACAAAGATTTCTATA 660
DB 601 GGTGGCAAGAAATAGAGCGCGGCGAGTACATTCCTTTCTATCAACAAAGATTTCTATA 660

RESULT 2
AAZ92760
ID AAZ92760 standard; cdna; 2112 Bp.
XX

QY 661 ACAAGACTTTTACTTAAACTTGCTAAGTTAGAGTTCAATTTGCTTTCAGTCATTGCAACAGG 720
DB 661 ACAAGACTTTTACTTAAACTTGCTAAGTTAGAGTTCAATTTGCTTTCAGTCATTGCAACAGG 720
QY 721 AAGAGCTCAGCCATGTGTCGAATGTGGAAGCTTTTCGATATCAAGAAAGAACGACCTT 780
DB 721 AAGAGCTCAGCCATGTGTCGAATGTGGAAGCTTTTCGATATCAAGAAAGAACGACCTT 780
QY 781 GTTTAAGAGATAGAAATTTGTAATGCTACTTTTGGGAGCTAGTTCAGGCTATGAGCCAC 840
DB 781 GTTTAAGAGATAGAAATTTGTAATGCTACTTTTGGGAGCTAGTTCAGGCTATGAGCCAC 840
QY 841 AGTATTCGGGCTAGAGTTTCTTCAAAAAGCTTGTCTGTTATTAACCTTTATAGATG 900
DB 841 AGTATTCGGGCTAGAGTTTCTTCAAAAAGCTTGTCTGTTATTAACCTTTATAGATG 900
QY 901 ACACCTTATGATCGTATGCTACTTATGAAGAACTTAAAGATCTTACTGAAGCTGTTGAAA 960
DB 901 ACACCTTATGATCGTATGCTACTTATGAAGAACTTAAAGATCTTACTGAAGCTGTTGAAA 960
QY 961 GGTGTCATTAATCATGCTTTAGACACACTTCCAGAAATACATGAAACCGATATACAAATAT 1020
DB 961 GGTGTCATTAATCATGCTTTAGACACACTTCCAGAAATACATGAAACCGATATACAAATAT 1020
QY 1021 TCATGGATACATACAGAAATGGAAGAAATTTCTTGCRAAGGAGGGAAGAACAGATCTAT 1080
DB 1021 TCATGGATACATACAGAAATGGAAGAAATTTCTTGCRAAGGAGGGAAGAACAGATCTAT 1080
QY 1081 TTAATCTCGGCAAGAAATTTGTGAAGAGTTTGTGAAGAGCTTGTGAAGAGCTTGTGAAGAG 1140
DB 1081 TTAATCTCGGCAAGAAATTTGTGAAGAGTTTGTGAAGAGCTTGTGAAGAGCTTGTGAAGAG 1140
QY 1141 GGGCAAAATGAGGAGACATACCAACCACTGAAGAGCATGATCCAGTTGTATCAATCTG 1200
DB 1141 GGGCAAAATGAGGAGACATACCAACCACTGAAGAGCATGATCCAGTTGTATCAATCTG 1200
QY 1201 GGGTGTCAACCTGCTTACCAACCTTGTATCTTGGCATGAGTGATATATTCACAAAG 1260
DB 1201 GGGTGTCAACCTGCTTACCAACCTTGTATCTTGGCATGAGTGATATATTCACAAAG 1260
QY 1261 AGTCTGTGGAATGGGCTGCTCTGCACTCTCTTTTAGATATCTCAGGTATATCTTGGTC 1320
DB 1261 AGTCTGTGGAATGGGCTGCTCTGCACTCTCTTTTAGATATCTCAGGTATATCTTGGTC 1320
QY 1321 GACGCTAAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCATCGA 1380
DB 1321 GACGCTAAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCATCGA 1380
QY 1381 GCGTTGAAGTTATATGAGGAATATATGTCATGAGGAGTATGCCCAACCTTGAATTT 1440
DB 1381 GCGTTGAAGTTATATGAGGAATATATGTCATGAGGAGTATGCCCAACCTTGAATTT 1440
QY 1441 ACAAGAAAGTAGAAGATGTGGAAGATATAAAACGAGAGTACCTCACAACTAAAAACA 1500
DB 1441 ACAAGAAAGTAGAAGATGTGGAAGATATAAAACGAGAGTACCTCACAACTAAAAACA 1500
QY 1501 TTCCAAGCCGCTTATGATGGCTGTGATCTATTTGTCAGAGTTTCTTGAAGTTCAATATG 1560
DB 1501 TTCCAAGCCGCTTATGATGGCTGTGATCTATTTGTCAGAGTTTCTTGAAGTTCAATATG 1560
QY 1561 CAGGAAAGGATTAACCTTCAACGATATGGGAGACGAATACAAACATCTCTATTAAGTCTCTAC 1620
DB 1561 CAGGAAAGGATTAACCTTCAACGATATGGGAGACGAATACAAACATCTCTATTAAGTCTCTAC 1620
QY 1621 TCGTTTATCTCTATGATATATGAGGATCC 1649
DB 1621 TCGTTTATCTCTATGATATATGAGGATCC 1649

amorphadiene to be produced and hence increasing artemisin production. It may then be economically feasible to extract artemisin from such plants for use as an antimalarial. The present sequence represents a positive CDNA clone encoding amorphadiene synthase identified in an *A. annua* CDNA library using a 538 bp amorphadiene synthase probe (AA92157) in an exemplification of the present invention

SA	Sequence	2112 BP	689 A	376 C	423 G	624 T	0 U	0 Other
QY	Query Match	99.5%	Score	1640.6	DB 3	Length	2112	
QY	Best Local Similarity	99.8%	Pred. No.	0				
QY	Matches 1643	Conservative	0	Mismatches	4	Indels	0	Gaps
QY	2	CATGGCACTTACAGAAGAAAAACCTATTTCGCGCCCAATTGCGCCAACTTTCTCCCTCCAAAGCAATTG	61					
DB	24	CATGTCACTTACAGAAGAAAAACCTATTTCGCGCCCAATTGCGCCAACTTTCTCCCTCCAAAGCAATTG	83					
QY	62	GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA	121					
DB	84	GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA	143					
QY	122	TGATTTAAAAAAGAGTGGCGCAACTACTATAAAGAAGCTTTGGATATTCTTATGAAACA	181					
DB	144	TGATTTAAAAAAGAGTGGCGCAACTACTATAAAGAAGCTTTGGATATTCTTATGAAACA	203					
QY	182	TGCCAAATTTGTTGAAGCTGATGTAGTAAATTCACAGCCTTGGAAATACCGTATCACCTTTGA	241					
DB	204	TGCCAAATTTGTTGAAGCTGATGTAGTAAATTCACAGCCTTGGAAATACCGTATCACCTTTGA	263					
QY	242	ACGGGAGATTGATCATGCAATTCGAATGTATTTATGAAACATATGGTGATAACTGGGAATGG	301					
DB	264	ACGGGAGATTGATCATGCAATTCGAATGTATTTATGAAACATATGGTGATAACTGGGAATGG	323					
QY	302	TGACCGCTCTTCCTTATGGTTCCTTATGCGAAAGCAAGATATTATGTTACATGTGA	361					
DB	324	TGACCGCTCTTCCTTATGGTTCCTTATGCGAAAGCAAGATATTATGTTACATGTGA	383					
QY	362	TGTTTTCAATAACTAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAAATGATCT	421					
DB	384	TGTTTTCAATAACTAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAAATGATCT	443					
QY	422	TGAAGGTTTGCTTGAGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATATT	481					
DB	444	TGAAGGTTTGCTTGAGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATATT	503					
QY	482	AGAAGATGCTCTTGTTTTTACAGCATCTCGTCTTAGCATATTAGCAAAAGATGCTTTTTTC	541					
DB	504	AGAAGATGCTCTTGTTTTTACAGCATCTCGTCTTAGCATATTAGCAAAAGATGCTTTTTTC	563					
QY	542	TACAAACCCCGCTCTTTTTTACCGAATACACGGGCATTAAGCAACCCCTTTGGAAAG	601					
DB	564	TACAAACCCCGCTCTTTTTTACCGAATACACGGGCATTAAGCAACCCCTTTGGAAAG	623					
QY	602	GTTGCCAAGAATAGAGGGGGCGGCGAGTACATCTCTTTCTATCAACCAAGATTCTCATAA	661					
DB	624	GTTGCCAAGAATAGAGGGGGCGGCGAGTACATCTCTTTCTATCAACCAAGATTCTCATAA	683					
QY	662	CAAGACTTTACTTAAACCTTGCTTAAGTTAGAGTTCAATTTGCTTCAGTCAATGGCAAGA	721					
DB	684	CAAGACTTTACTTAAACCTTGCTTAAGTTAGAGTTCAATTTGCTTCAGTCAATGGCAAGA	743					
QY	722	AGAGCTCAGGCATGTGTCGAATGTGGAAAGCTTTTCGATATCAAGAAGAACGCACCTTG	781					
DB	744	AGAGCTCAGGCATGTGTCGAATGTGGAAAGCTTTTCGATATCAAGAAGAACGCACCTTG	803					
QY	782	TTTAAGAGATAGAATTGTTGAATGCTACTTTTTGGGGACTAGGTTTCAGGCTATGAGCCACA	841					
DB	804	TTTAAGAGATAGAATTGTTGAATGCTACTTTTTGGGGACTAGGTTTCAGGCTATGAGCCACA	863					
QY	842	GTAATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTATAAATCTTTATAGATGA	901					
DB	864	GTAATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTATAAATCTTTATAGATGA	923					

QY	902	CAC	TATGATCGTATGTTACTTATGAAACCTTAAAGTCTTTACTGAAGCTTTGAAAG	961
Db	924	CAC	TATGATCGTATGTTACTTATGAAACCTTAAAGTCTTTACTGAAGCTTTGAAAG	983
QY	962	GTG	CTCAATTACATCTTACACACACCTTCCAGAAATACATGAACCGATATACAAATATT	1021
Db	984	GTG	CTCAATTACATCTTACACACACCTTCCAGAAATACATGAACCGATATACAAATATT	1043
QY	1022	CAT	GATACATACACAGAAATGGAAGAAATTTCTTGCAAAAGGAGGGAAGAACAGATCTATT	1081
Db	1044	CAT	GATACATACACAGAAATGGAAGAAATTTCTTGCAAAAGGAGGGAAGAACAGATCTATT	1103
QY	1082	TAA	CTGGCGCAAAAGAAATTTGTTGAAGAGTTTGTAGAAAACCTGATGGTTGAAGCAAAATG	1141
Db	1104	TAA	CTGGCGCAAAAGAAATTTGTTGAAGAGTTTGTAGAAAACCTGATGGTTGAAGCAAAATG	1163
QY	1142	GGC	AAATGAGGACACATACCAACCACTGAAGAGATGATCCAGTTGTATATCACTACTCG	1201
Db	1164	GGC	AAATGAGGACACATACCAACCACTGAAGAGATGATCCAGTTGTATATCACTACTCG	1223
QY	1202	CGG	TCTAACCTGTACAAACAACTTGTATCTTGGCATGAGTGATATATTCACAAAAGA	1261
Db	1224	CGG	TCTAACCTGTGTACAAACAACTTGTATCTTGGCATGAGTGATATATTCACAAAAGA	1283
QY	1262	GTCT	GTGGAATGGGTCTCTGACCTCTCTTTTATAGATCTACAGGTATCTTGGTCTG	1321
Db	1284	GTCT	GTGGAATGGGTCTCTGACCTCTCTTTTATAGATCTACAGGTATCTTGGTCTG	1343
QY	1322	ACG	CTAATGATCTCATGACCCACAGCCGACGCAAGCAAGAAACATAGTTCATCGAG	1381
Db	1344	ACG	CTAATGATCTCATGACCCACAGCCGACGCAAGCAAGAAACATAGTTCATCGAG	1403
QY	1382	CCTT	GAAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCAAACCTTGATTTA	1441
Db	1404	CCTT	GAAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCAAACCTTGATTTA	1463
QY	1442	CAAG	GAATGAGATGTGTGGAAGATATAAACCAGAGTACCTCACAACTAAAACAT	1501
Db	1464	CAAG	GAATGAGATGTGTGGAAGATATAAACCAGAGTACCTCACAACTAAAACAT	1523
QY	1502	TCC	AGGCGGTATTTGATGGCTGTGATCTATTGTGCCAGTTCTTGAAGTTCAATATGC	1561
Db	1524	TCC	AGGCGGTATTTGATGGCTGTGATCTATTGTGCCAGTTCTTGAAGTTCAATATGC	1583
QY	1562	AGG	AAAGGATTAATTTACACAGTATGGGAGACGAATACAAACATCTCTAAAGTCTCTACT	1621
Db	1584	AGG	AAAGGATTAATTTACACAGTATGGGAGACGAATACAAACATCTCTAAAGTCTCTACT	1643
QY	1622	CGT	TATCCTATGATATATGAGGATC	1648
Db	1644	CGT	TATCCTATGATATATGAGTACC	1670

RESULT 3
AAZ92763
ID AAZ92763 standard; DNA; 2076 BP.

XX	AC	AAZ92763;
XX	DT	19-JUN-2000 (first entry)
XX	DE	Artemisia annua amorphadiene synthase DNA clone.
XX	DE	Artemisia annua amorphadiene synthase; FPP;
XX	DE	Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;
XX	DE	farnesyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;
XX	DE	sesquiterpene; flavouring; fragrance; ss.
XX	OS	Artemisia annua.
XX	PH	Key Location/Qualifiers
XX	PH	6..1646
XX	PH	/tag= a
XX	PH	/product= "Artemisia annua amorphadiene synthase"

XX	EP982404-A1.
XX	01-MAR-2000.
XX	27-AUG-1998; 98EP-00202854.
XX	27-AUG-1998; 98EP-00202854.
XX	(WALL/) WALLAART T E.
XX	Wallaart TED, Bouwmeester HJ;
XX	WPI; 2000-258617/23.
XX	P-PSDB; AAY81272.
XX	New isolated DNA sequences and polypeptides comprising amorpha-4,11-diene synthase activity, useful for production of amorphadiene and/or artemisinin.
XX	Disclosure; Page 10-14; 41pp; English.
XX	The invention relates to DNA encoding Artemisia annua amorpha-4,11-diene synthase (AAZ92754). This enzyme catalyses the cyclisation of farnesyl pyrophosphate (FPP), a ubiquitous precursor of primary and secondary metabolites, into amorpha-4,11-diene. Amorphadiene is a precursor of artemisin, a sesquiterpene lactone endoperoxide produced by A. annua which is a promising antimalarial drug candidate. Amorphadiene synthase is a branch point enzyme in the artemisin biosynthetic pathway, and the cyclisation of FPP is therefore likely to be the rate-limiting step.
XX	Artemisin is present in very low concentrations in A. annua, making it expensive for use as an antimalarial drug, and attempts at organic synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are also useful as flavour and fragrance compounds in the food and perfume industries. Terpenes also play a role in plant-insect interactions, such as the attraction or repulsion of insects by plants. In addition, dihydroartemisinin acid, an intermediate in the metabolic route from amorphadiene to artemisinin in A. annua, can be used as an antioxidant.
XX	DNA encoding amorphadiene synthase may be used to generate transgenic plants able to synthesise amorphadiene synthase. Preferably, the plants used for amorphadiene production are plants which naturally produce sesquiterpenes, as these plants already have the basic sesquiterpene synthetic pathways and storage compartments. A. annua may be transformed with amorphadiene synthase expression constructs, enabling more amorphadiene to be produced and hence increasing artemisin production. It may then be economically feasible to extract artemisin from such plants for use as an antimalarial. The present sequence represents a DNA clone encoding A. annua amorphadiene synthase, which is not referred to further in the specification
XX	Sequence 2076 BP; 681 A; 363 C; 412 G; 620 T; 0 U; 0 Other;

Query Match 99.3%; Score 1637.4; DB 3; Length 2076;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	2	CATGGC	ACTTACAGAGAAAAACCTATTCGCCCATTCGCCAATTTCTCTCAAGCATTTG	61
Db	5	CATGTC	ACTTACAGAGAAAAACCTATTCGCCCATTCGCCAATTTCTCTCAAGCATTTG	64
QY	62	GGGAGAT	CAGTTCTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTCAA	121
Db	65	GGGAGAT	CAGTTCTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTCAA	124
QY	122	TGATTTAAAA	AAAAAGAGTGGCGCACTACTATAAAGAGAGCTTTGGATATTCCTATGAACA	181
Db	125	TGATTTAAAA	AAAAAGAGTGGCGCACTACTATAAAGAGAGCTTTGGATATTCCTATGAACA	184
QY	182	TGCCAATTTG	TTCGAAGCTGATTGATGAAATTCACGCGCTTGGAAATACCGTATCACTTTGA	241
Db	185	TGCCAATTTG	TTCGAAGCTGATTGATGAAATTCACGCGCTTGGAAATACCGTATCACTTTGA	244
QY	242	ACGGGAGAT	GATGATCATGCAATTCGAATGTATTTATGAACAACATATGGTGATAACTCGAATGG	301

245	Db		ACGGG	GATTGATCATG	ATGCAATG	CGAATGATAT	TATGA	AACAATATGGT	GATAACTGGAATGG	304	
302	Qy		TGACCG	CTCTT	CCCTATG	TTGGTCCG	TTATG	CGAAAGCAAGG	ATATTTATGTTACATGTGA	361	
305	Db		TGACCG	CTCTT	CCCTATG	TTGGTCCG	TTATG	CGAAAGCAAGG	ATATTTATGTTACATGTGA	364	
362	Qy		TGTTTT	CAATPA	ACTATATA	AAGACAAAA	TGGAG	CGCTTCAAG	CAATCGTTAGCTAATGATGT	421	
365	Db		TGTTTT	CAATPA	ACTATATA	AAGACAAAA	TGGAG	CGCTTCAAG	CAATCGTTAGCTAATGATGT	424	
422	Qy		TGAGGT	TTTCTT	GAGTTGTAC	GNAGCA	CTTCTAT	GAGGGTACCT	GGGAGATATATTT	481	
425	Db		TGAGGT	TTTCTT	GAGTTGTAC	GNAGCA	CTTCTAT	GAGGGTACCT	GGGAGATATATATTT	484	
482	Qy		AGAAG	TGCTCT	TGGTTT	TACAGAT	CTCGTCTT	AGCATTATG	ACAAAAAGATGCTTTTTC	541	
485	Db		AGAAG	TGCTCT	TGGTTT	TACAGAT	CTCGTCTT	AGCATTATG	ACAAAAAGATGCTTTTTC	544	
542	Qy		TACAAA	CCCCCG	CTCTTTT	TACGGAAT	PACAA	CGGGCACTAAAG	CAACCCCTTTGGAAAAAG	601	
545	Db		TACAAA	CCCCCG	CTCTTTT	TACGGAAT	PACAA	CGGGCACTAAAG	CAACCCCTTTGGAAAAAG	604	
602	Qy		GTTTGC	CAAGAA	TAGAGG	CGGCGAG	TACATTC	CTTNTCA	CAACAAAGATTCTCATAA	661	
605	Db		GTTTGC	CAAGAA	TAGAGG	CGGCGAG	TACATTC	CTTNTCA	CAACAAAGATTCTCATAA	664	
662	Qy		CAAGAC	TTTTACT	TAAACT	TGCTAAG	TTAGAG	TTCAATTTG	CTTCAGTCATTGCACAAAGGA	721	
665	Db		CAAGAC	TTTTACT	TAAACT	TGCTAAG	TTAGAG	TTCAATTTG	CTTCAGTCATTGCACAAAGGA	724	
722	Qy		AGAGCT	CAGC	CAATGTGT	GCAAA	TGGTGG	AAAGCTTTCG	ATATCAAGAAAGAACGCACCTTG	781	
725	Db		AGAGCT	CAGC	CAATGTGT	GCAAA	TGGTGG	AAAGCTTTCG	ATATCAAGAAAGAACGCACCTTG	784	
782	Qy		TTTAA	GAGATAG	AAATTTG	TGAATG	CTACTTTT	TGGGACTAG	GTTCAGGCTATGAGCCACA	841	
785	Db		TTTAA	GAGATAG	AAATTTG	TGAATG	CTACTTTT	TGGGACTAG	GTTCAGGCTATGAGCCACA	844	
842	Qy		GTATTC	CCGGCT	TAGAGT	TTTTCTT	CAAA	AAAGCTGTG	CTTATTA	ACTCTTATAGATGA	901
845	Db		GTATTC	CCGGCT	TAGAGT	TTTTCTT	CAAA	AAAGCTGTG	CTTATTA	ACTCTTATAGATGA	904
902	Qy		CAC	TTATGATG	CGGTATG	GTACTT	ATGAAG	AACTTAAG	ATCTTTACT	CAAGCTGTGGAAG	961
905	Db		CAC	TTATGATG	CGGTATG	GTACTT	ATGAAG	AACTTAAG	ATCTTTACT	CAAGCTGTGGAAG	964
962	Qy		GTGGT	CAATTA	CATGCTT	AGACAC	CACTTC	AGAAATAC	ATGAAACCGG	ATATACAAATPATTT	1021
965	Db		GTGGT	CAATTA	CATGCTT	AGACAC	CACTTC	AGAAATAC	ATGAAACCGG	ATATACAAATPATTT	1024
1022	Qy		CATGG	ATACATAC	ACAGAA	ATGGA	AGATTTCTT	GTGAAAGAGG	GGAGACAGATCTATT	1081	
1025	Db		CATGG	ATACATAC	ACAGAA	ATGGA	AGATTTCTT	GTGAAAGAGG	GGAGACAGATCTATT	1084	
1082	Qy		TAACT	CGGGCA	AGAAATTTG	TG	TGAAAGAG	TTTGTGTAG	AAACCTGATGGT	TGGAAGCAAAATG	1141
1085	Db		TAACT	CGGGCA	AGAAATTTG	TG	TGAAAGAG	TTTGTGTAG	AAACCTGATGGT	TGGAAGCAAAATG	1144
1142	Qy		GGCAAA	TGAGGG	ACACATAC	CAAC	CACTGTGA	AGCATGATCC	AGTTCAGTTCATTTACTCGG	1201	
1145	Db		GGCAAA	TGAGGG	ACACATAC	CAAC	CACTGTGA	AGCATGATCC	AGTTCAGTTCATTTACTCGG	1204	
1202	Qy		CGGTG	CTAAC	CTGTTAC	ACAA	CACTTGTTAT	CTTGG	CATGAGTGATATATTTCA	CAAAGA	1261
1205	Db		CGGTG	CTAAC	CTGTTAC	ACAA	CACTTGTTAT	CTTGG	CATGAGTGATATATTTCA	CAAAGA	1264
1262	Qy		GTCTG	TGGAAT	GGGCTG	CTCTG	CACCTCTCT	TTTTTTAG	ATACTCAGGTATACTTTGGTGC	1321	
1265	Db		GTCTG	TGGAAT	GGGCTG	CTCTG	CACCTCTCT	TTTTTTAG	ATACTCAGGTATACTTTGGTGC	1324	
1322	Qy		ACGCCT	AAAAATG	ATCTCT	ATGAC	CCCAAG	CGCGAGCA	AGAAAGAAAA	CATAGTTTCATCGAG	1381

Db	1325	ACGCCTAAATGATCTCATGACCCCAAGGCCGAGCAAGAAAGAAAACAATAGTTTCATCGAG	1388
Qy	1382	CTTTGAAAGATTATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAAACCTTGATTTA	1441
Db	1385	CTTTGAAAGATTATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAAACCTTGATTTA	1444
Qy	1442	CAAGGAAAGTAGAAGATGTGTGGAAAGATATAAACCGAGAGTACCTCACAACCTAAAAACAT	1501
Db	1445	CAAGGAAAGTAGAAGATGTGTGGAAAGATATAAACCGAGAGTACCTCACAACCTAAAAACAT	1504
Qy	1502	TCCAAAGCCGTTATTGATCGCTGTGATCTATTGTGCGCAGTTTCTTTGAAAGTTCAATATGC	1561
Db	1505	TCCAAAGCCGTTATTGATCGCTGTGATCTATTGTGCGCAGTTTCTTTGAAAGTTCAATATGC	1564
Qy	1562	AGGAAAGGATAACTTTCACACGTATGGGAGACCAATACAAACATCTCTATAAAGTCTCTACT	1621
Db	1565	AGGAAAGGATAACTTTCACACGTATGGGAGACCAATACAAACATCTCTATAAAGTCTCTACT	1624
Qy	1622	CGTTTATCTTATGAGTATATGAGGATC	1648
Db	1625	CGTTTATCTTATGAGTATATGAGTACC	1651
RESULT 4			
AAZ92757			
ID AAZ92757 standard; cDNA; 538 BP.			
XX	AAZ92757;		
AC			
XX	19-JUN-2000 (first entry)		
DT			
XX	Artemisia annua amorphadiene synthase 538 bp cDNA probe.		
DE			
XX	Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;		
KW	farnesyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;		
KW	sesquiterpene; flavouring; fragrance; hybridisation probe; ss.		
XX			
OS	Artemisia annua.		
XX			
PN	EP982404-A1.		
XX			
PD	01-MAR-2000.		
XX			
PF	27-AUG-1998; 98BP-00202854.		
XX			
PR	27-AUG-1998; 98BP-00202854.		
XX			
PA	(WALL/) WALLAART T E.		
XX			
PI	Wallaart TED, Bouwmeester HJ;		
XX			
DR	WPI; 2000-258617/23.		
DR	P-PSDB; AAY81720.		
XX			
PT	New isolated DNA sequences and polypeptides comprising amorpha-4,11-diene		
PT	synthase activity, useful for production of amorphadiene and/or		
PT	artemisinin.		
XX			
PS	Example 2; Fig 7; 4lpp; English.		
XX			
CC	The invention relates to DNA encoding Artemisia annua amorpha-4,11-diene		
CC	synthase (AAZ92754). This enzyme catalyses the cyclisation of farnesyl		
CC	pyrophosphate (FPP), a ubiquitous precursor of primary and secondary		
CC	metabolites, into amorpha-4,11-diene. Amorphadiene is a precursor of		
CC	artemisin, a sesquiterpene lactone endoperoxide produced by A. annua		
CC	which is a promising antimalarial drug candidate. Amorphadiene synthase		
CC	is a branch point enzyme in the artemisin biosynthetic pathway, and the		
CC	cyclisation of FPP is therefore likely to be the rate-limiting step.		
CC	Artemisin is present in very low concentrations in A. annua, making it		
CC	expensive for use as an antimalarial drug, and attempts at organic		
CC	synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are		
CC	also useful as flavour and fragrance compounds in the food and perfume		
CC	industries. Terpenes also play a role in plant-insect interactions, such		

CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
CC nerolidol, and monoterpenoid. This polynucleotide sequence represents the
CC DNA encoding an H64 protein used in the terpenoid biosynthesis method of
CC the invention. NOTE: This sequence is not shown in the specification. It
CC has been obtained from electronic data supplied with this specification
CC from the European Patent Office.

XX
SQ Sequence 1973 BP; 612 A; 361 C; 434 G; 566 T; 0 U; 0 Other;

Query Match 24.0%; Score 396.2; DB 9; Length 1973;
Best Local Similarity 53.8%; Pred. No. 4.4e-96;
Matches 877; Conservative 0; Mismatches 733; Indels 21; Gaps 2;

Qy	22	AACCTATTGCGCCCATTCGCCAATTCCTCCCAAGCAATTTGGGGAGATCAGATTTCTCATCT	81
Db	53	AAGTGTGTCGGGCACAGCAAAATTTAAACCTAGCGTTTGGGGAGATCGTTTGTCTAAT	112
Qy	82	ATCAAAAGCAAGTAGAGCAAGGGGTGGACAGATAGTGAATTTAAAAAAGAGTGC	141
Db	113	ATGCCGAAGACATTATAACTCAAACTCAAAATCAAGAACAAAGTTGAGGAGCTGAACAAG	172
Qy	142	GGCAACTACTAAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTGAGGCTGA	201
Db	173	TAGTGAGGAAGAGTATTTCATTAATGCTGCTGATGATTCCTCAATCAACTGAAGCTAA	232
Qy	202	TTGATGAAATTCACGCCCTTGGAAATACCGTATCACTTTTGAACGGGAGATTGATCATGCAT	261
Db	233	TTGATGAATCCAGCGCCTCGGTGGCTTACCATTTCGAAGCGAATAGATCAAGCCC	292
Qy	262	TGCAATGATTTTGAACAATATGATGTGATTAACCTGGAATGGTG-----ACCGCTCTT	312
Db	293	TGGAACGATATACATGAGACATATCAAGATATTCATGATGGTGGTGATCTGTACAATGTTG	352
Qy	313	CTTATGTTTCGGCTTTATGCGAAGCAAGATATTTGTTACATGTGATGTTTCAATA	372
Db	353	CTCTTCGTTTTCGGGTACTCAGCGGACATGGATATAATGTTTCCGCGATGATTCAACA	412
Qy	373	ACTATAAGACAAAAATGGAGGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGC	432
Db	413	AGTTCAAGATACATATGTTGACTACAGAAAGCTTGCTACTGATCTTTCTGGTATGC	472
Qy	433	TTGAGTTGTAAGAGCAACTTCTATGAGGTTACCTGGGGAGATTTATTTAGAAGATGCTC	492
Db	473	TGAGCTTTTATGAGCGGCCCATCTGAGGGTGCATGGAGAAAAATTACTTGAAGAGGCTC	532
Qy	493	TTGCTTTTACAGATCTGCTTACATATGACAAAGATGCTTTTCTACAAACCCCG	552
Db	533	TGGTTTTTACCACCACTCATCTCC-----AGTCAGCAAGTGCANAAAGCTCTT	580
Qy	553	CTCTTTTACCAGAAATACAAAGCGGCACTAAAGCAACCCCTTTTGGAAAAGGTTGCCAAGAA	612
Db	581	TGCTGAACAACAATAACTGNAAGCGGTAGAGAGACCTCTACTAAAACATATGGAGAGT	640
Qy	613	TAGAGGGCGGCGACATTCCTTTCTATCAACAACAAGATTTCTCATACAAGACTTTAC	672
Db	641	TAGGTGTCGGCGTTACATGTCAATATATCAAGATGAAGCTTCAATACAGTGAATAATAC	700
Qy	673	TTAACTTCTAAGTTAGATTCATTTGCTTCAGTTCATTTGACAGGAGAGCTCAGCC	732
Db	701	TGAAACTTGCAAAATTTGATTTTAATGTTTGTTCAGTGTTCACAAAAGGAACTCAGTG	760
Qy	733	ATGTGTGCAATTTGGGAAAGCTTTTCGATATCAAGAAAGCAACGCTCTTTTAAAGATA	792
Db	761	ACATTTCAAGATGGTACAGGAACCTGGACTTTTGCAGAGGAGTGCCTTTTGTCTCGATA	820
Qy	793	GAATTTGTAATGCTACTTTTGGGACCTAGGTTTCAGGCTATGAGCCACAGTATCCCGGG	852
Db	821	GGATCGTGGAGTTGTTCTTTTGGATAGCAGGAATATTTTCGAACCTGTAATAGCTTTTG	880
Qy	853	CTAGAGTTTCTTCAAAAAGCTGTGCTGTTTATTAATCTTTATAGATGACACTTATGATG	912
Db	881	GGAGACACATTTCTGACTAACTGATTTGAGATACACAGTAAATGGATGATATGATGATG	940

Qy	913	CGTATGCTACTTATGAAGAACTTAAAGATCTTTACTGAAGCTGTTGAAAGGTGGTCAATTA	972
Db	941	CATTGCGTACATTCGAAGAACTCGTCACTTCTGCTGAAGCAATTCGACAGTGGGATGCAA	1000
Qy	973	CATGCTTAGACACATTCGAGATACATGAACCGATATACAAATTTATTCATGGATACAT	1032
Db	1001	GTTGCATGGATCAACTGCCAGCTATATGCAACCATTTTATATACACATCTTCGATGTTA	1060
Qy	1033	ACACAGAAATGGAAGATTTCTTCAAGGAGGGAAGAACAGATCTATTTAACTCGCGCA	1092
Db	1061	TCGATGAAGTTGAAGGAGCTGCAAGGAGGAGATCTTCCGAATTCATCTAGCACA	1120
Qy	1093	AGAATTTTGAAGAGTGTGTTTGAAGAACTGATGGTGAAGCAAAATTTGGGCAAAATGAGG	1152
Db	1121	AGAATTTTGAAGAACTCAAGCCAGGCTCTACTTCGCTGAGGCCATATGTTTCCACGAAG	1180
Qy	1153	GACACATACCAACCTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAAC	1212
Db	1181	GATGCCACCCCAAAATGGATGGGTATATGCGAGTTGCGGCATCTTCTGCGGTAAACCA	1240
Qy	1213	TGCTTACCAACAACTTGTATCTTGGCATGATGATATTTCAAAAAGAGTCTGTGCGAAT	1272
Db	1241	TGCTTTCCGTCGTCTTTTAGTAGGCTAGGAGATTAACCAAAATTTGNAATTCGAGT	1300
Qy	1273	GGCTGTCTCTGCACTCTCTTTTATAGATCTCAGGTATATCTTGGTTCGACGCCCTAAATG	1332
Db	1301	GGCTGACCAATGAGCTTAAATCCCTTAGAGCTTCGAATACCATATTTAGGCTTATGGATG	1360
Qy	1333	ATCTCATGCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCCTTGAAGTT	1392
Db	1361	ACATTCGCTGGTCAAGTTTGAGAAAGAGAGAGGCGCATGTTGCTTCTAGTATTTGATGCT	1420
Qy	1393	ATATGAAGGAATATATGTCAATGAGGAGTATGCCCAACCTTGATTTACAAGGAAGTAG	1452
Db	1421	ACATGAATGAATACGGGTTTTCAGAGCAAGAGACATTTGATATCTTCAACAAACGAATTG	1480
Qy	1453	AAGATGTGTGGAAGATATAAACCGAGAGTACCTCAACATTAATAAAACATTCGAAGCCGT	1512
Db	1481	TGGATTCGTGGAAGATATAAACGAAGAGTTTCTGAGACCCACTGCTGCTCCAGTCCCTG	1540
Qy	1513	TATTGATGGCTGATCTATTTTGGCCAGTTTCTTGAAGTTCAATATCGCAGAAAGGATA	1572
Db	1541	TGCTTAATCGTGTCTTAACTAACCCGAGTGGTGTGATCTGCTTTACAAAAGGGGAGATG	1600
Qy	1573	ACTTCACACTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCTTA	1632
Db	1601	CTTTCAGCAATGTCGGAATACTGATGAAGATTTGTTGCTGCAATGTTTATGATCCAG	1660
Qy	1633	TGAGTATATGA 1643	
Db	1661	TGCCACTCTGA 1671	

RESULT 6
ADE82704
ID ADE82704 standard; DNA; 2605 BP.

XX ADE82704;
AC ADE82704;
XX 29-JAN-2004 (first entry)
XX Terpenoid biosynthesis related H64 strawberry DNA #15.

DE isoprenoid; bio-active compound synthesis; pesticide; dermatological;
KW cytosolic; immunosuppressive; virucide; flavour; fragrance;
KW bio-control agent; food additive; food industry; pest control;
KW degreasing solvent; plasticizer; dye carrier; dental caries;
KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
KW anti-retroviral; monoterpene alcohol linalool;
KW sesquiterpene alcohol nerolidol; monoterpene; strawberry; gene; de.
XX Fragaria x ananassa.
XX

PN EP1231273-A1.
XX
PD 14-AUG-2002.
XX
PF 12-FEB-2001; 2001EP-00200488.
XX
PR 12-FEB-2001; 2001EP-00200488.
XX
PA (PLAN-) PLANT RES INT BV.
XX
PI Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;
XX WPI; 2003-879727/82.
XX
XX Novel recombinant nucleic acid encoding proteinaceous molecule, useful
PT for producing flavor, fragrance and/or biocontrol agent which is useful
PT as food additive in processed food industry and as antimicrobial agent.
XX
XX Disclosure; Page; 52pp; English.
XX
XX The invention relates to a novel isolated or recombinant nucleic acid or
CC its functional fragment, encoding a proteinaceous molecule essentially
CC capable of isoprenoid bio-active compound synthesis when provided with a
CC suitable substrate under appropriate reaction conditions. The novel
CC recombinant isoprenoid bio-active compound synthesis nucleic acid and its
CC protein have the following activities: pesticide, dermatological,
CC cytostatic, immunosuppressive, and virucide. The novel recombinant
CC isoprenoid bio-active compound synthesis nucleic acid is useful for
CC producing flavour, fragrance, and/or a bio-control agent, by transforming
CC or transfecting a suitable host with the recombinant isoprenoid bio-
CC active compound synthesis nucleic acid, expressing the recombinant
CC isoprenoid bio-active compound synthesis nucleic acid in the presence of
CC a suitable substrate, and optionally isolating the formed product. The
CC bio-control agent is useful as an anti-microbial agent, as a food
CC additive in the processed food industry to modify the taste of syrups,
CC ice-creams, frozen desserts, yogurts, confectionery and like products, as
CC a flavouring agent for oral medications and vitamins, and for providing
CC additional flavour/aroma in beverages including alcoholic beverages. The
CC bio-control agent is also useful for enhancing or reducing flavour,
CC aroma, fragrance or scent of plants, natural products, and/or synthetic
CC or artificial products, and for the industrial synthesis of nature
CC identical flavour/aroma substances and/or artificial flavour/aroma
CC substances. The bio-control agent is also useful as a pest control agent
CC for the biological control of the interaction between plants and insects
CC and/or plants and microorganisms, for providing flavour/aroma in
CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
CC products, personal care products and health care products, as a
CC disinfectant additive and in the preparation of a composition. The novel
CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
CC fragments is useful as a molecular marker or diagnostic tool. The protein
CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
CC acid is useful for the production of an antagonist e.g. an antibody or
CC its functional equivalent which is useful for inhibiting the synthesis of
CC the bio-control agent. A composition, containing the bio-control agent,
CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
CC the aroma and/or taste of food or non-food products, and/or protection of
CC food or non-food products against fungal contamination and/or pest
CC infestation. The composition is also useful for the biological control of
CC pests, for the protection of stored products and for the prevention or
CC treatment of disease. The bio-control agent is useful as a degreasing
CC solvent, plasticizer and dye carrier. The composition is useful for
CC replacing potentially carcinogenic synthetic food additives currently
CC used. The composition is also useful for treating dental caries, dental
CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and
CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
CC compound synthesis nucleic acid or its protein is useful for the
CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
CC nerolidol, and monoterpene. This polynucleotide sequence represents the
CC DNA encoding an H4 protein used in the terpene biosynthesis method of
CC the invention. NOTE: This sequence is not shown in the specification. It
CC has been obtained from electronic data supplied with this specification
XX from the European Patent Office.

SQ	Sequence	2605 BP; 821 A; 454 C; 541 G; 789 T; 0 U; 0 Other;
Query Match	22.4%;	Score 369.8; DB 9; Length 2605;
Best Local Similarity	53.6%;	Pred. No. 6.3e-89;
Matches	875;	Conservative 0; Mismatches 732; Indels 26; Gaps 4;
Qy	22	AACCTATTGCGCCCATTCGCCAACTTTCTCCAAAGCATTTGGGGAGATCAGTTTCTCATCT 81
Db	744	AAGTTGTTGGCGGCACACAAATTTTAAACCTAGCGTTGGGGAGATCGTTTGCTAACT 803
Qy	82	ATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAAAAGAAAGTGC 141
Db	804	ATGCGGAAGACATTATTAATCAAACTCAAAATGCAAGAACAAAGTTGAGGAGCTGAAAACAAG 863
Qy	142	GGCAACTACTAAAAGAGACTTTGGATATTCCTATGAACATGCCCAATTTGTTGGAAGCTGA 201
Db	864	TAGTGAGGAGGAAGATATTCAATATGCTGCTGATGATTTCTTCATCATCAACTGAAGCTAA 923
Qy	202	TTGATGAAATTCAAACGCGCTTGGAAATACCGTATCACTTTGAACGGGAGATTTGATCATGCAT 261
Db	924	TTGATGAAATCCAGGGCTCGGTGGCTTACCATTTCGAAGCGGAATAGATCAAGCCC 983
Qy	262	TGCAATGTATTATGAAACATATATGTGTGATACTGGAATGGTG-----ACCGTCTTT 312
Db	984	TGGAACGTATATCATGAGACATATCAAGATATTTCATGATGGTGGTGTATCTGTACAATGTTG 1043
Qy	313	CTTTATGGTTCGCTTATGCGAAGCAAGGATATATGTTACATGTGATGTTTCAATA 372
Db	1044	CTCTTCGCTTTTCGGCTACTCAGGCGACATGATATAATGTTTTCCTCGCATGATTTCAACA 1103
Qy	373	ACTATAAGACAAAAATGGAGCGTTCAAGCAATCCTTAGCTAATGATGTTGAAGGTTTG 432
Db	1104	AGTTCAAGATACTAATGGTGACTACAGAAAGGCTTGGTCACTGATCTTTCTGGTATGC 1163
Qy	433	TTGAGTTGTACGAAGCAACTTCTTATAGGGTACCTGGGGAGATTATATTAGAAGATGCTC 492
Db	1164	TGAGCTTTTATAGGCGGCCCATCTGAGGGTGCTAGGAGAAAAATTTACTTGAAGAGGCTC 1223
Qy	493	TTGCTTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCG 552
Db	1224	TGGTTTTTACCACCACCTCATCTCC-----AGTCAGCAAGTGCAAAAGAGCTCTT 1271
Qy	553	CTCTTTTACCAGAAATCAACGGGCATCAAGCAACCCCTTTGGAAAGGTTTGCACCAAGAA 612
Db	1272	TGCTGAAAACAAATTAAGTGAAGCGGTGAGAGACTACT---AAAACATATGGAGAGGT 1328
Qy	613	TAGAGGCGCGCAGTACATTCCTTTCTATCAACAAAGATTTCTATACAAAGACTTTAC 672
Db	1329	TAGTGCTCGCGGTTACATGTCAATATATCAAGATGAAGCTTCATACAGTGAATAATTAC 1388
Qy	673	TTAACTTGCTAAGTTAGATTCATTTGCTTCAGTCAATGCGCAAGGAGAGCTCA--G 730
Db	1389	TGAAACTTGCAAAATTAGATTTTAAATGTTGTTTCAGTGTTTACACAAAAAGGAACTCAGTG 1448
Qy	731	CCATGTGTCAAAATGGTGAAGGCTTTTCGATATCAAGAAAGACGCACTTTGTTTAAGAGA 790
Db	1449	ACATTCCTAGATGGTTCACAGGAAGCTGAGCTTTTGCAGAGGAGTGCCTTTTGTCTGAGA 1508
Qy	791	TAGAATTTGTAATGCTACTTTTGGGGACTAGGTTTCAGGCTATGAGCCACAGTATTTCCCG 850
Db	1509	TAGGATCGTGAGTTGTTCTTTTGGATAGCAGGAATATATTTGGAACCTGAAATACGCTTT 1568
Qy	851	GGCTAGAGTTTCTTCACAAAGGCTGTTGCTGTATTAATCTTTATAGATGACACTTATGA 910
Db	1569	TAGGAGACACTTCTGACTAACTGATTGAGATAACCAAGTAATGGAATGATGTATGATGA 1628
Qy	911	TGCGTATGGTACTTATGAGAACTTAAGATCTTTTACTGAAGCTGTTTGAAGGTGGTCAAT 970
Db	1629	TGCATTCGGTACATTCGAGAACTCGTCTCACTTGCTGAAGCAATTCAGAGTGGATGC 1688
Qy	971	TACATGCTTAGACACACTTCCAGAAATACATGAACCGGATATACAAATTTATCATGATAC 1030
Db	1689	AAGTTGATGATCAACTGCCAGACTATATGCAACCACTTTTATATTACACTTCTGTGATGT 1748

```
QY 1031 ATACACGAATGAAGAAATTTCTTCAAGAGGAGGAGAAAGACAGATCTATTTAACTGCGG 1090
Db 1749 TATCGATGAAGTGAAGAGGAGCTGACAAAGCAAGAGAGATCTTACCGAATTCACATCGC 1808
QY 1091 CAAGAATTTGTGAAGAGTTTGTAGAAA CCGTATGTTGTGAAGCAAAATGGCGCAATGA 1150
Db 1809 AAAAGAAATATGAAGAAATCAAGCAGGCTCTACTTCGCTGAGGCGAGATGTTCCACGA 1868
QY 1151 GGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCAATCTACTGGCGTGCTAA 1210
Db 1869 AGGATGACCCCAAAATATGATGATATATGCGAGTTGCGGCACTCTTCTGCGTAACAC 1928
QY 1211 CCTGTTACAAACAACCTGTTATCTTTGGCATGATATATTCACAAAAGAGTCTGTGCA 1270
Db 1929 CATGCTTTCCTGCTGCTTTAGTAGGATGGAGACATTAACAAATTTGAATTCGA 1988
QY 1271 ATGGGCTGCTCTGCACTCTCTCTTTTGTAGATCTCAGGTATCTTGGTGGAGCGCTAAA 1330
Db 1989 GTGGCTGACCAATGAGCCTAAATCTTAGAGCTTCGAATACCATATTTAGGCTTATGGA 2048
QY 1331 TGATCTCATGCCACACGCGCGACGACAGAAAGAAACATAGTTCTATCGAGCTTGAAG 1390
Db 2049 TGACATCTCGGTGACAAAGTTTGAAGAAAGAGAGAGGCGCATGTTGCTTCAAGTATTGATG 2108
QY 1391 TTATATGAAGAAATATAATGTCAATGAGGAGTATGCCAAACCTTGATTTACAAGGAAT 1450
Db 2109 CTACATGATGAATACGGGTTTCAGACCAAGACAAATGATATCTTCAACAAACGAAT 2168
QY 1451 AGAAGATGTGGAAAGATATAAACCAGAGAGTACCTCAACAATTAATAAATCCAAAGGC 1510
Db 2169 TGTGATTCGTGGAAGGATATAAACAAGAGTTCTGAGACCCACTGCTGCCAGTCCC 2228
QY 1511 GTTATGATGGCTGTAATATTTTGTGCGAGTTTCTTGAAGTTCATATGCGAGGAAGA 1570
Db 2229 TGTGCTTAATCGTGTCTTAACCTAACCCGAGTGTGATCTGCTTTCAAAAGGGGAGA 2288
QY 1571 TAACTTCACAGTATGGGAGAGCAATACAAACATCTCAATAAGTCTCTACTCGTTTATCC 1630
Db 2289 TGCCTTCAGCATGTCGGAACATGATGAAGATTGTATGTGCTGCAATGTTATTGATCC 2348
QY 1631 TATGATATATGA 1643
Db 2349 AGTGCCACTCTGA 2361
```

RESULT 7

```
AA89865
ID AA89865 standard; cdna; 1879 BP.
XX
AC AA89865;
XX
DT 08-OCT-1999 (first entry)
XX
DE L. esculentum germacrene C synthase cdna (insert pLE 20.3).
XX
KW Germacrene C synthase; tomato; transgenic; pathogen resistance;
ZW flavour alteration; neutral value; ss.
XX
OS Lycopersicon esculentum.
XX
FH Key Location/Qualifiers
CDS 39..1685
FT /*tag= a
FT /*product= "germacrene C synthase"
XX
FN WO9938957-A1.
XX
PD 05-AUG-1999.
XX
PF 02-FEB-1999; 99WO-US002133.
XX
PR 02-FEB-1998; 98US-0073579P.
```

```
XX
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
XX
PI Colby SM, Crook JE, Lemaux PG, Croteau RB;
XX
XX WPI; 1999-479180/40.
DR P-PSDB; AAY27062.
XX
```

PT New Germacrene C synthase protein and nucleic acid useful for producing
transgenic plants with increased resistance to pathogens.

XX
PS Claim 12; Fig 6; 60pp; English.

XX This represents the nucleotide sequence of the cDNA insert pLE 20.3
encoding a Lycopersicon esculentum germacrene C synthase protein. The
germacrene C synthase gene is useful for generating transgenic plants
with an increased host resistance to pathogens including bacteria, rusts,
herbivores, microbes and fungi. Recombinant nucleic acid comprising a
specific expression operably to the germacrene synthase gene allows tissue
specific expression of the protein to deter e.g. leaf-eating herbivores,
and allows production of plants with an altered flavour and odour
profiles to deter e.g. fruit-eating herbivores, or to enhance the plants
attractiveness for pollinators, or increase the nutraceutical value of
the plant. Polypeptide fragments of the protein are useful as immunogens
to raise antibodies, and the polynucleotides are useful as probes or
primers for detecting, cloning and confirming alleles and homologues of
the germacrene synthase gene

XX SQ Sequence 1879 BP; 665 A; 312 C; 362 G; 540 T; 0 U; 0 Other;

Query Match 22.1%; Score 364; DB 2; Length 1879;
Best Local Similarity 52.9%; Pred. No. 2e-87;
Matches 865; Conservative 0; Mismatches 745; Indels 24; Gaps 3;

```
QY 29 TCGCCCATTCGCCAATCTTCTCCAGCATTTGGGAGATCAGTTTCTCATCTCAAA 88
Db 65 TCGCCCATTCGCCAATCTTCTCCAGCATTTGGGAGATCAGTTTCTCATCTCA 124
QY 89 GCAGTAGAGCAAGGGTGGGAACAGATAGTGAATGATTTAAAGAAAGAGTCCGCACT 148
Db 125 TGAATTAATCAAGAAAGTTGAAGTTGATGATGATGATGATGATGATGATGAT 184
QY 149 ACTAAAAAGAGCTTTGGATATTCCTATGAAACATGCCAATTTGTTGAAGCTGATTGA 208
Db 185 GCTGGTGAACCTTCGCACATAGCCTCAAA-----GCTTGTGTTGATAGACGC 235
QY 209 AATTCAACGCTTGGAAATACCGTATACCTTTGAACGGAGATGATGATGATGATGATG 268
Db 236 GATGCAACGATTTGGGAGTGGCTTATCATTTCCGATAATGAAATTTGAAACATCCATT 295
QY 269 TATTTATGAACAT-----ATGGTGAATGATGATGATGATGATGATGATGATGAT 316
Db 296 CATTTTGTGATGATCGTCCAAACAGAAATGATGATGATGATGATGATGATGATGAT 355
QY 317 ATGCTTCGCTTATGCGAAAGCAAGGATATTTATGTTACATGATGATGATGATGATGAT 376
Db 356 TCGTTTTCGACTGTGAGGCAACAGGCCATTAATGATGATGATGATGATGATGATGAT 415
QY 377 TAAAGACAAAATGAGCGTTTCAAGCAATCGTTAGTAAATGATGATGATGATGATGATG 436
Db 416 CACCAACCAAGATGGGAAATTTCAAGGAAACACTTACTAATGATGATGATGATGATGAT 475
QY 437 GTTGTAGAGCAACTTCTATGAGGTTACCTGGGAGATTAATATAGAGAGATGCTCTGG 496
Db 476 TTTGTATGAAGCATCAATCTGAGAGTGGTAATGAGGAGATTTCTTGAAGAGCTCTTAC 535
QY 497 TTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTTACAAAACCCGCT 556
Db 536 ATTTACCAACCTCATCTCGAGTCTTATGCTCCCACTTGAGCAATAATAAATCTCTCT 595
QY 557 TTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGATA 616
```

596 TAAGTTGAAGTTGGTGAAGCCCTTAAGTCAAGCTTATTCGATGACTTTACCAAGGATGGG 655
617 GCGCGCGAGTACATTCCTCTTATCAACAAAGATTCATCAAGACTTTTACTTAA 676
656 AGCTAGAAAATACATATCCATTTACGAAAACAATGATGCACACCACTTTGCTTTTGA 715
677 ACTTGTAAAGTTAGAGTTCAATTTGCTTCAGTCAATTCACAAGGAAGCTCAGCCATGT 736
716 ATTGTCTAAATGGATTTTAACTGCTGCAAAAGTTTCAAAAAGAGCTTAGTATCT 775
737 GTGCAATGCTGAAAGCTTTTCGATATCAAGAAGACGCACTTGTTTTAAAGAGATAGAAT 796
776 TACAAGGCTGGAAGATTTTGATTTTGCAATAAATATCCATATCAAGAGACAGTT 835
797 TGTGTAATGCTACTTTTGGGAGTAGGTCAGGCTATGAGCAAGATTTCCCGGGTAG 856
836 GGTGTAGTGTACTTCTGGATATTAGAGTGTATTTTGGAGCAAAATATATGTCGTGGAG 895
857 AGTTTCTTCCACAAAAGCTGTGCTGTATTAACCTTTATAGATGACACTTATGATGGTA 916
896 AAAATGATGACAAAAGTACTCAACTGACCTTCAATTTATGACACACTTTTGATGCTTA 955
917 TGTACTTATGAAGAACTTAAAGATCTTTACTGAAGCTGTTGAAAGGTGGTCAATTACATG 976
956 TGCAACCTTTGACGAAGCTTGTGAGCTTTCAATGATGCAATCCAGAGATGGGATGCTATGC 1015
977 CTTAGACACTTCCAGATACATGAAACCGATATACAAATTTATTCATGATACATACAC 1036
1016 AATTGATTCATTAACAACCATATATGACCTGCTTATCAAGCTCTTCTAGACATTTACAG 1075
1037 AGAAATGGGAAGATTTCTTCAAGAGGAGGAAGACAGATCTATTTAACTGCGGCAAGA 1096
1076 TGAATGGAAACAAGTGTGTCAAAGAGGTAACTGGACCGGTATATCTATGCAAAAA 1135
1097 ATTGTGAAAGAGTTTGTAGAAACCTGTATGGTTGAAGCAAAATGGGCAAAATGA --GGG 1153
1136 TGAGATGAAAGAGTTGTGAGAGCTTATTTAAGGAAACCAATGGTTGATGATTTGA 1195
1154 ACATATACCAACCACTGAGAGATGATCAAGTTGTAATCATTTACTGCGGTGCTACCT 1213
1196 CCATATTTCCAAAATATGAGGAACAAGTGGAGATGCAATCGTAGTCTGCTATATGAT 1255
1214 GCTTACAAACCTGTATCTTGGATGAGTATATTCACAAAAGAGCTGCGAATG 1273
1256 GATATCAAAACCTGTGTGCTGATAGAGAAATTTATATCCACAGACTTTTGAATG 1315
1274 GGTGTCTCTGCACCTCCTCTTTTATAGATCTCAGGTATATCTTGGTGGAGCCCTTAAATGA 1333
1316 GTTGATGATGATGCTGTGATGCTTGGCTTCCGATTCGTCAGAGCAATGAACGA 1375
1334 TCTCATGCCCAAGCCGAGCAAGAAAGAAATATGATTCATGAGCCCTTGAAAGTTA 1393
1376 TATGTTGGACATGAAGATGAACAAGAAAGAGACATGTAGCTTCACTTATTGAATGTTA 1435
1394 TATGAAGGAATAAATGTCATGAGGAGTATGCCAAACCTTGATTTTACAGGAAGTGA 1453
1436 CATGAAGATTTGAGCTTCAAGCAAGAGACTTACATTAAGTTCTGTAAGAGGTTCAC 1495
1454 AGATGTGTGGAAGATATAACCGAGAGTACCTCAACAACCTTAAACAACTTCCAGGCGCTT 1513
1496 CAATGATGGAAGGACATAAACAACATTTCTCCGTCCAACTGAAGTACCAATGTTGT 1555
1514 ATTGATGGCTGTGATCTATTTGTCGAGTTTCTTGAAGTTCAATATGCAAGAAAGATAA 1573
1556 CTTGAACGAGTTCTAAATTTGACACGCTGTGGCTGACACGCTTATATAAGAGAAAGATAC 1615
1574 CTTACACGATGGGAGACCAATACAAACATCTCATAAAGTCTCTACTCTCTTATCTCTAT 1633
1616 ATATCAAGCCCAAGGAAACCTTAAACATGATTTATTCATATTAATGATCTGT 1675
1634 GAGTATATGAGGAT 1647
1676 CAAAATATAAATAT 1689

RESULT 8

ADE82701

ID ADE82701 standard; DNA; 1672 BP.

XX ADE82701;

XX 29-JAN-2004 (first entry)

XX Terpenoid biosynthesis related H64 strawberry DNA #14.

XX isoprenoid; bio-active compound synthesis; pesticide; dermatological;
XX cytosolic; immunosuppressive; virucide; flavour; fragrance;
XX bio-control agent; food additive; food industry; pest control;
XX degrading solvent; plasticizer; dye carrier; dental caries;
XX dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
XX anti-retroviral; monoterpene alcohol linalool;
XX sesquiterpene alcohol nerolidol; monoterpene; strawberry; gene; ds.

Fragaria x ananassa.

EPI231273-A1.

14-AUG-2002.

12-FEB-2001; 2001EP-00200488.

12-FEB-2001; 2001EP-00200488.

(PLAN-) PLANT RES INT BV.

Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;

WPI; 2003-879727/82.

Novel recombinant nucleic acid encoding proteinaceous molecule, useful
for producing flavor, fragrance and/or biocontrol agent which is useful
as food additive in processed food industry and as antimicrobial agent.

Disclosure; Page; 52pp; English.

The invention relates to a novel isolated or recombinant nucleic acid or
its functional fragment, encoding a proteinaceous molecule essentially
capable of isoprenoid bio-active compound synthesis when provided with a
suitable substrate under appropriate reaction conditions. The novel
recombinant isoprenoid bio-active compound synthesis nucleic acid and its
protein have the following activities: pesticide, dermatological,
cytostatic, immunosuppressive, and virucide. The novel recombinant
isoprenoid bio-active compound synthesis nucleic acid is useful for
producing flavour, fragrance, and/or a bio-control agent, by transforming
or transfecting a suitable host with the recombinant isoprenoid bio-
active compound synthesis nucleic acid, expressing the recombinant
isoprenoid bio-active compound synthesis nucleic acid in the presence of
a suitable substrate, and optionally isolating the formed product. The
bio-control agent is useful as an anti-microbial agent, as a food
additive in the processed food industry to modify the taste of syrups,
ice-creams, frozen desserts, yogurts, confectionery and like products, as
a flavouring agent for oral medications and vitamins, and for providing
additional flavour/aroma in beverages, including alcoholic beverages. The
bio-control agent is also useful for enhancing or reducing flavour,
aroma, fragrance or scent of plants, natural products, and/or synthetic
or artificial products, and for the industrial synthesis of nature
identical flavour/aroma substances and/or artificial flavour/aroma
substances. The bio-control agent is also useful as a pest control agent
for the biological control of the interaction between plants and insects
and/or plants and microorganisms, for providing flavour/aroma in
cosmetics, creams, sun-protectant products, hair conditioners, cleaning
products, personal care products and health care products, as a
disinfectant additive and in the preparation of a composition. The novel
recombinant isoprenoid bio-active compound synthesis nucleic acid or its
fragments is useful as a molecular marker or diagnostic tool. The protein
of the novel recombinant isoprenoid bio-active compound synthesis nucleic

CC acid is useful for the production of an antagonist e.g. an antibody or
CC its functional equivalent which is useful for inhibiting the synthesis of
CC the bio-control agent. A composition, containing the bio-control agent,
CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
CC the aroma and/or taste of food or non-food products, and/or protection of
CC food or non-food products against fungal contamination and/or pest
CC infestation. The composition is also useful for the biological control of
CC pests, for the protection of stored products and for the prevention or
CC treatment of disease. The bio-control agent is useful as a degrading
CC solvent, plasticizer and dye carrier. The composition is useful for
CC replacing potentially carcinogenic synthetic food additives currently
CC used. The composition is also useful for treating dental caries, dental
CC plaque and skin disorders, and for immunosuppressive, anti-leukemia and
CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
CC compound synthesis nucleic acid or its protein is useful for the
CC synthesis of monoterpenes alcohol linalool and sesquiterpene alcohol
CC nerolidol, and monoterpenoid. This polynucleotide sequence represents the
CC DNA encoding an H64 protein used in the terpenoid biosynthesis method of
CC the invention. NOTE: This sequence is not shown in the specification. It
CC has been obtained from electronic data supplied with this specification
CC from the European Patent Office.

XX SQ Sequence 1672 BP; 514 A; 315 C; 385 G; 458 T; 0 U; 0 Other;

Query Match 22.0%; Score 363.6; DB 9; Length 1672;
Best Local Similarity 54.2%; Pred. No. 2.5e-87;
Matches 887; Conservative 0; Mismatches 714; Indels 35; Gaps 6;

QY 22 AACCTATTGCGCCCATTCGCCAATTTCTCCCAAGCATTTGGGGAGATCAGTTTCTCATCT 81
DB 53 AAGTTGTTGGCGGCACAGCAAAATTTAAACCTAGCGTTTGGGGAGATCGGTTTGTCACT 112
QY 82 ATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAAGAAAGTGC 141
DB 113 AT-----GCCGAAGACATTTAACTCAAACTCAAAATGCAAGCAAAAGTTGAGGAGCTGA 166
QY 142 GGCAACTACTAAAAGAGCTTTGGATATCTCTATGAACA---TGCCAAATTTGTTGAAGC 198
DB 167 AACANGTAGAGGAAGATATCTAATGCTGCTGATGATTTCTCACTCAACTGAAGC 226
QY 199 TGATTGATGAATTCAGCGCTTGGAAATACCGTATCACTTTGAACGGGAGATTGATCATG 258
DB 227 CAATTGATGAATCCAGCGCTCGGTGGCTTACCAATTTGAAAGCGAATAGATCAAG 286
QY 259 CATTGCAATGATTTATGAACAATATGTTGATGATACTGGAATGGTG-----ACCGCT 309
DB 287 CCTGGAAACGTTATACATGAGACATATCAAGATATTCATGATGGTGGTGTATCTGTAATG 346
QY 310 CTTCCTTATGCTTCGCTTATGCGAAGCAAGGATATTTATGATGATGATGATTTTCA 369
DB 347 TTGCTCTGCTTTTCGGCTACTCAGGCGACATGGGATATAATGTTTCTCGGATGATTCA 406
QY 370 ATAACTATAAGACAAATAATGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTT 429
DB 407 ACAAGTTCAAGATACATAATGCTGACATACAGAAAGCTTGGTCACTGATCTTCTGGTA 466
QY 430 TGCTTGAGTTGTAAGAAAGCAATCTTATGAGGTTACCTGGGAGATTATATPAGAAGATG 489
DB 467 TGCTGAGCTTTATGAGCGGCCCATCTGAGGTTGATGAGGAAAAATTTACTTGAAGAGG 526
QY 490 CTCTTGCTTTTACAGATCTGCTTAGCATATGACAAAGATGCTTTTCTACAAACC 549
DB 527 CTCTGGTTTTTACCACCACTCATCTCC-----AGTCAGCAAGTGCAAAAGCT 574
QY 550 CCGCTCTTTTTTACCAGAAATACAAAGCGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAA 609
DB 575 CTTTGTCTGAAACACAAATACTAGAGCGGTAGAGAGACTACT---AAAAGTATGGAGA 631
QY 610 GAATAGAGGCGCGGAGPACATTCCTTTCTATCAACAAAGATTTCTCATACAGAGATT 669
DB 632 GGTTAGGTGCTCGCGGTTTACATGTCAATATATCAAGATGAAGCTTCAATACAGTGAATTT 691
QY 670 TACTTAACTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCACTTGCACAAAGAGAGCTCA 729

DB 692 TACTGAAACTTGCAAAATTAGATTTTAAATGTTGTTCAAGTGTATTACACAAAAGGAACCTCA 751
QY 730 --GCCATGTGTGCAAAATGGTGGAAAGCTTTGCATATCAAGAAAGAACGCACTTTGTTTAAAG 787
DB 752 GTGACATTCCTTAAGATGGTACAAAGGAACCTGGACTTTGCAAGGAGGATGCTTTTGTCTCG 811
QY 788 AGATAGAAATTTGTTGAATGCTACTTTTGGGAGCTAGGTTTCAGGCTATGAGCCACAGTATTC 847
DB 812 AGATAGGATCGTGGAGTTGTTCTTTGGATAGCAGGAATATATTTCGAACCTGGAATACCT 871
QY 848 CCGGGCTAGAGTTTCTTTCACAAAAGCTGTTGCTGTTTATACTCTTATAGATGACACTTA 907
DB 872 CTTTAGAGACACATCTCTGACTAACTGATGATGATGATGATGATGATGATGATGATGAT 931
QY 908 TGATGCTATGGTACTTATGAAAGAACTTAAAGATCTTTTCTGAAAGCTGTTGAAAGGTGTC 967
DB 932 TGATGATTCGGTACATTCGAAGAACTCGTCAACTGCTGACTGAAAGCAATTTGACAGGTGGA 991
QY 968 AATTACATGCTTAGACACACTTCCAGAAATACATGAAACCGGATATACAAATTTATTCATGGA 1027
DB 992 TGCAAGTTGTCATGGATCAACTGCCAGACTATATGCAACCACTTTTATATTACACTTCTGGA 1051
QY 1028 TACATACACAGAAATGGAAGAATTTCTTGCAAGGAGGGAAGAACAGATCTATTATAACTG 1087
DB 1052 TGTATTCGATGAAGTTGAAGAGGAGCTGACAAAGCAAGGAAGATCTTACCGAATTCACCTA 1111
QY 1088 CGCAAAAGAAATTTGTGAAAGAGTTTGTAGAAACCTGATGTTGAAAGCAAAATGGGCAAA 1147
DB 1112 CGCAAAAGAAATTTATGAAGAAATCAAGCCAGCTCTACTTCCGCTGAGGCCAGATGTTCCA 1171
QY 1148 TGAGGACACATACCAACCACTGAAAGAGCATGATCCAGTTGTTGATCATCTACTTGGCGGTGC 1207
DB 1172 CGAAGGATGACCCCAAAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
QY 1208 TACCTGCTTTACAAACAACTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1267
DB 1232 CACCATGCTTTCGTCGTGCTTTTAGTAGCATGGGAGACATTAACAAAATTTGAAAT 1291
QY 1268 CGAATGGCTGCTCTGCACTCTCTCTTTTAGATATCTAGGTATCTTGGTTCGACGCT 1327
DB 1292 CGAGTGGCTGACCAATGAGCTTAAATCCCTTAGAGCTTCGAATACCATATTTAGGCTTAT 1351
QY 1328 AATGATCTCATGCCCAAGCCGAGCAAGCAAGAAAGAAACATAGTTTCATCGAGCTTGA 1387
DB 1352 GGATGACATTTGCTGGGTACAAAGTTTGAGAAAGAGAGAGGAGGATGTTGCTTCAAGTATTGA 1411
QY 1388 AAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCAAACCTTGAATTTTCAAGGA 1447
DB 1412 TTGCTACATGAATGAATACCGGGTTTTCAGAGCAAGAGACAAATGATATCTTCAACAAAG 1471
QY 1448 AGTAGAAGATGTTGGAAAGATATAAACCCGAGAGTACCTCACAACTAAAAACATTTCCAAAG 1507
DB 1472 AATTGTGGATTCGTGGAAGGATATAAACGAAAGAGTTTCTGAGACCCACTGCTGCTCCAGT 1531
QY 1508 GCGGTTATGATGGCTGATGATCTATTGTCGCCAGTTTCTTGAAGTTCAATATGAGGAAA 1567
DB 1532 CCCTGTGCTTAATCGTGTCTTAAACCTTAACCCAGAGTGGTTGATCTGCTTTTCAAAAGGGG 1591
QY 1568 GGATTAATTCACACGATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTA 1627
DB 1592 AGATGCTTTCACGCATGTCGAAACATGATGAAAGATTTGATTTGCTGCAATGTTTATTGA 1651
QY 1628 TCCTATGATATATGA 1643
DB 1652 TCCAGTGCCACTCTGA 1667

RESULT 9
AAx89866
ID AAx89866 standard; cdna; 2024 BP.
XX
AC AAx89866;

XX 08-OCT-1999 (first entry)

XX DE L. esculentum germacrene C synthase cDNA (insert pLE 14.2).

XX DE Germacrene C synthase; tomato; transgenic; pathogen resistance;

XX KW flavour alteration; neuroreputical value; ss.

XX OS Lycopersicon esculentum.

XX Key Location/Qualifiers

XX FH 32..1678

XX CDS /*tag= a

XX /product= "germacrene C synthase"

XX PN MO9938957-A1.

XX PD 05-AUG-1999.

XX PF 02-FEB-1999; 99WO-US002133.

XX PR 02-FEB-1998; 98US-0073579P.

XX PA (UNIW) UNIV WASHINGTON STATE RES FOUND.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Colby SM, Crock JE, Lemaux PG, Croteau RB;

XX DR WPI; 1999-479180/40.

XX DR P-PSDB; AAY27063.

XX PT New Germacrene C synthase protein and nucleic acid useful for producing

XX PT transgenic plants with increased resistance to pathogens.

XX PS Example; Fig 8A-B; 60pp; English.

XX CC This represents the nucleotide sequence of the cDNA insert pLE 14.2

XX CC encoding a Lycopersicon esculentum germacrene C synthase protein. The

XX CC germacrene C synthase gene (cDNA insert pLE 20.3) is useful for

XX CC generating transgenic plants with an increased host resistance to

XX CC pathogens including bacteria, rusts, herbivores, microbes and fungi.

XX CC Recombinant nucleic acid comprising a promoter sequence operably to the

XX CC germacrene synthase gene allows tissue specific expression of the protein.

XX CC to deter e.g. leaf-eating herbivores, and allows production of plants

XX CC with an altered flavour and odour profiles to deter e.g. fruit-eating

XX CC herbivores, or to enhance the plants attractiveness for pollinators, or

XX CC increase the neuroreputical value of the plant. Polypeptide fragments of

XX CC the protein are useful as immunogens to raise antibodies, and the

XX CC polynucleotides are useful as probes or primers for detecting, cloning

XX CC and confirming alleles and homologues of the germacrene synthase gene

XX SQ Sequence 204 BP; 705 A; 332 C; 379 G; 608 T; 0 U; 0 Other;

Query Match 21.7%; Score 357.6; DB 2; Length 2024;

Best Local Similarity 52.7%; Pred. No. 1.1e-85;

Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

Qy 29 TCGGCCCATTCGCAATCTTCCTCCAGCATTTGGGAGATCAGTTTCCTCATCTATCAAAA 88

Db 58 TCGGCCCTGGCTAATTTTCAACCATCTGTTGGGATATCATTTCTTCTTATACTCA 117

Qy 89 GCAAGTAGCAAGGGTGGACAGATAGTGAATTTTAAAAAAGAGTGGGCACT 148

Db 118 TGAATTTACTAATCAAGAAAAAGTTGAAGTTGATGATGATCAAGAGACAATTAGAAAAAT 177

Qy 149 ACTAAAAAGAGCTTTGGATATTTCTTATGAACATGCCAATTTGTTGAAGCTGATTGATGA 208

Db 178 GCTGGTGAACCTTCGACATAGACATCAAAA-----GCTTGTGTTGATAGACGC 228

Qy 209 AATTCACGCTTGGAAATACCGTATCATTTGAAAGGGAGATGATCATGCAATGGCAATG 268

Db 229 GATGCAACGATTTGGAGTGGCTTATCATTTTCGATAATGAATTGAACATCATTCATAA 288

Qy 269 TATTATGAACAT-----ATGGTATAACTGGATGGTACCGCTCTTCCTT 316

Db 289 CATTTTGTGTCATCGTCCAAACAGAAATGATAATGACAAACACCTTTAGTGTGTCTCT 348

Qy 317 ATGTTCCGCTTTATGCGAAAGCAAGGATATTTGTTACATGTCATGTTTCAATAAATA 376

Db 349 TCGTTTTCGATTTGTGAGGCAACAGGCCATTAATGTCCTTCAGATGTCCTTCAAGCAAT 408

Qy 377 TAAAGACAAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGTCTGA 436

Db 409 CACCAACCAAGATGGGAAATTTCAAGGAAACACTTACTAATGATGTCCTCAAGGATATTGAG 468

Qy 437 GTTGTAGAGCAACTTCTATGAGGTCACCTGGGAGATTTATATTAGAGATGCTCTTGG 496

Db 469 TTTGTATGAAGCATCACATCTGAGAGTGGGTAATGAGGAGATTTCTTGAAGAGCTCTTAC 528

Qy 497 TTTTACACGATCTCGTCTTAGCATTTATGACAAAAAGATGCTTTTCTACAAACCCGCTCT 556

Db 529 ATTACACCACTCATCTCGAGTCTATTGTCCTCAACTTGGAGCAATATATTAATCTCT 588

Qy 557 TTTTACCGAAATACACCGGCACCTAAAGCAACCCCTTTTGGAAAAAGTTGCCAAGAAATA 616

Db 589 TAAGTTTGAAGTTGGTGAAGCCTTAACTCAGCTTATTCGATGACTTTTACCAAGGATGG 648

Qy 617 GCGGGCGCAGTACATCTCTTCTATCAACACAGATTTCTCATACAGACTTTTACTTAA 676

Db 649 AGCTAGAAAAATCATATATCCATTTACGAAAAACAATGATGCAACCACTTTGCTTTTGA 708

Qy 677 ACTTGCTTAAGTTAGAGTTCAATTTGCTTCACTGATGCAAGGAAGAGCTCAGGCATCT 736

Db 709 ATTTGCTTAATTTGATTTTAACTGCTGCAAAAGTTTCACCAAGAGAGCTTAGTATCT 768

Qy 737 GTCCAAATGTTGAAAGCTTTTCGATATCAAGAAAGCAACCTTGTGTTTGAAGATAGAA 796

Db 769 TACAAGGTGTTGAAAGATTTGGATTTTGCAAAATAAATATCCATATGCAAGAGACAGTT 828

Qy 797 TGTGTAATGCTACTTTTGGGAGCTAGCTTCAGGCTATGAGCCACAGATTTCCGGGCTAG 856

Db 829 GGTGTAGTGTACTCTTGGATATTTAGGAGTGTATTTTGGAGCCAAAATAATAGTCGCGAG 888

Qy 857 AGTTTCTTCCAAAAAGCTCTGCTGTATAAATCTTTATAGATGACACTTATGATGCTGA 916

Db 889 AAAAATGATGACAAAAAGTACTCAACCTGACCTCATTTATGACGACACTTTTGTATGCTTA 948

Qy 917 TGTGCTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAAAGGTGGTCAATTTACATG 976

Db 949 TGCAACCTTTGAGCAACTTGTGATCTTTCAATGATGCAATCCAGAGATGGGATGCTTAATGC 1008

Qy 977 CTTTAGACACTTCCAGAAATACATGAAACCGATATACAAATTATTCATGGATACATACAC 1036

Db 1009 AATTGATTTCAATACACCATATATGAGACTGCTTTATCAAGCTCTTCTAGACATTTACAG 1068

Qy 1037 AGAAATGGAGAAATTTCTTCAAGAGGAGGAAACAGATCTATTTAACTCGCGCAAGA 1096

Db 1069 TGAATGGAAACAGTGTGTTGCCAAAGAGGTAAACTGGACCGTGTATCTATGCAAAAAA 1128

Qy 1097 ATTTGTGAAGAGTTTGTGTAGAAACCTGATGTTGAGAGCAAAATGGGCAATGA---GGG 1153

Db 1129 TGAGATGAAAAAGTTGGTGTAGAGCTTATTTTAAAGGAAACCAATGGTTGAATGATGTA 1188

Qy 1154 ACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGCGGTGCTAACT 1213

Db 1189 CCATTTCCAAATATGAGGAACAAGTGGAGAAATGCAATCGTAAGTGTGCTATATGAT 1248

Qy 1214 GCTTACAAACACTTGTATCTTGGCATGAGTGAATATTTCAAAAAAGAGTCTGTGCAATG 1273

Db 1249 GATATCAACAACTTGTGTGTCGTAATGAAGAAATTTATATCCACAGAGACTTTTGAATG 1308

Qy 1274 GCGTGTCTGCACTCTCTTTTATAGTACTAGGATATCTTGTGTCAGCGCTTAATGA 1333

Db 1309 GTTGATGAATGAGTCTGTGATTTGTCGAGCTTCGCACTTGGTTCAGAGCAATGAACGA 1368

Qy 1334 TCTCATGACCCCAAGGCGCGAGCAAGAAAGAAACATAGTTTCTCAGAGCTTTGAAAGTTA 1393

769 TACAAGGTGGTGAAGATTGGATTTCCTCAATAAATATCCATATGCAAGACAGAGTT 828
797 TGTGAATGCTACTTTGGGAGCTAGGTTGAGCTATGACGACAGTATTCCTGGGCTAG 856
829 GGTGAGTGTACTTCTGGATATTAGGAGTGTATTTGAGCCAAATATAGTCGCGAG 888
857 AGTTTCTTCAAAAGCTGTGCTGTATTAATCTCTATAGATGACACTTATGATGCGTA 916
889 AAAATGATGACAAAAGTACTCAACCTGACCTCAATTTATGAGACACTTTTGTATGCTTA 948
917 TGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTGTAAGGTGTGTCATTAATGATG 976
949 TGAACCTTTGACGAACCTTGTGACCTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
977 CTTAGACACACTTCAGATACATGAAACCGATATACAAATTTATCATGATACATACAC 1036
1009 AATTGATTCATACCAACCATATATGAGACCTCTTATCAAGCTCTCTAGACATTTTACAG 1068
1037 AGAATGGAAGAAATTTCTTGAAGAGGAGGAAGACAGATCTATTAACTGCGGCAAGA 1096
1069 TGAATGGAACAGTGTGTCGAAGAGGTAACTGGACCTGTATCTATGTCGCAAAA 1128
1097 ATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGA ---GGG 1153
1129 TGAGATGAAAAGTTGGTGAGAGCTATTTTAAAGGAACCCCAATGGTTGAATGATGTCGA 1188
1154 ACATATCAACCACTGAAGAGCATGATCCAGTTGTAATCATCTGCGCGGTGCTAACCT 1213
1189 CCATATTCCTCAATATGAGGACCAATGAGATGCAATCGAAGTGTGCTGCTATATGAT 1248
1214 GCTTACAACTCTGTTTCTTGGATGATGATATATTCACAAAGAGCTCTGTGCAATG 1273
1249 GATATCAACACTCTGTTGTCGATGATGATGATATATCCACGACATTTTGAATG 1308
1274 GGCTGTCTCTCACTCTCTTTTATGATATCTAGGATATCTGTTGTCGACGCTAAATGA 1333
1309 GTTGATGAATGATGCTGTGATTTGTCGAGCTTCGACATTTGATTCGACAGCAATGAACGA 1368
1334 TCTCATGACCAACAGCGCGGACGAGGAGAAAGATAGTTTATCGAGCTTGAAGTTA 1393
1369 TATTGTTGACATGAAGATGACACGAGGAGATGATGCTTCACTTATTGAATGTTA 1428
1394 TATGAAGAAATATATGTCATGAGGAGTATGCCAAACCTTGATTTACAAGGAGTAGA 1453
1429 CATGAAGATATGAGCTTCAAGCAGAGACTTACATTAAGTCTTCAAGAGGTCAC 1488
1454 AGATGTGGAAGATATAAACCGAGGAGTACCTCAACATAAACAATTCGAAGCGGTT 1513
1489 CAATGATGGAAGGACATAAACAACAAATCTCCGTCACAACTGAAGTACCAATGTTGT 1548
1514 ATTGATGCTGTGATCTATTGTCGAGTTTCTTGAAGTTCAATATGCGAGGAGATAA 1573
1549 CTTGAACGAGTCTTAAATTTGACACGCTGTGCTGACAGCTTATATGAAGAGAAATAC 1608
1574 CTTACACGTATGGGAGACGAATACAAACATCTCTATAAAGTCTCTACTGTTTATCCTAT 1633
1609 ATATTCACCGCAAGGAAACTTAAACATGATTAATCAATACTTAATGATCTGT 1668
1634 GAGTATATGAGAT 1647
1669 CAAAATATAAAT 1682

RESULT 11
AAA38928

ID AAA38928 standard; DNA; 1944 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

25-AUG-2000 (first entry)

Potato vetaspiradiene synthase DNA sequence SEQ ID NO:31.

Synthase; protein co-ordinate data; active site; modification; terpenoid;
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
isoprenoid; breeding programme; fragrance; flavour; pheromone;
defensive agent; pigment; antitumor; steroid hormone;
signal transduction pathway; bile acid; affinity purification;
photoreceptor; enzymatic synthesis; nutrient supplement;
immunological reagent; ds.
Solanum tuberosum.

WO200017327-A2.

30-MAR-2000.

17-SEP-1999; 99WO-US021419.

18-SEP-1998; 98US-0100993P.

22-APR-1999; 99US-0130628P.

23-AUG-1999; 99US-0150262P.

(KENT) UNIV KENTUCKY RES DEPT.

(SALK) SALK INST BIOLOGICAL STUDIES.

Chappell J, Manna KR, Noel JP, Starks CW;

WPI; 2000-292839/25.

P-PSDB; AAY90843.

Novel terpene synthase enzymes, useful for producing terpene

hydrocarbons, e.g. fragrances or antitumor agents, are derived from known

enzymes by specific amino acid alterations.

Disclosure; Page 392-395; 450pp; English.

The present invention describes an isolated terpene synthase (I)

comprising a region with at least 20% identity to region 265-535 of a 548

amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha

-carbon atoms (alphaC) that have interatomic distances, between each

other, within tabulated ranges, have a centre point (within a sphere of

radius 2.3 Angstrom) within tabulated ranges, and have an ordered

arrangement of R groups (defining aa side chains), excluding specific

tabulated arrangements (tables given in the specification). (I), and

related enzymes, are used to produce a wide range of terpenoids (e.g.

cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,

flavours, pheromones, defensive agents, pigments, antitumor agents,

components of signal transduction pathways, precursors of steroid

hormones and bile acids, as photoreceptors and as co-factor side chains.

Some synthases with little or no catalytic activity (and nucleic acids

encoding them) are used as controls in the analysis of products formed by

enzymatic synthesis; as nutrient supplements; for affinity purification

of isoprenoids; or to develop immunological reagents or nucleic acids for

monitoring expression of terpene synthase or inheritance of the gene in

plant breeding programs. The new synthases may produce novel terpene

products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent

sequences used in the exemplification of the present invention

XX Sequence 1944 BP; 604 A; 359 C; 398 G; 583 T; 0 U; 0 Other;

Query Match 20.0%; Score 330.4; DB 3; Length 1944;

Best Local Similarity 52.0%; Pred. No. 2.3e-78;

Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAACCTATTGCCCCATTGCCAATCTTCTCCCAAGCATTTGGGGAGATCAGTT 73

Db 92 AGAGGAGGAGATTGTTGCCCCCACTAGCTGATCTTCTCCAAAGTCTTTGGGGTATGCTTT 151

QY 74 TCTCATCTATCAAAAGCAAGTAGACGAGGGGTGGAAACAGATAGTGAATGATTTAAAAA 133

Db 152 TCATTCAATCTCCCTCGACATCAGATTGCTGGAAAAATATGCTCAAGAGATCGAAACTTT 211

QY 134 AGAAGTGGGCAACTACTAAAGAGCTTTGGATATCTCTATGAACATGCCAATTTGTT 193

Db 212 GAAGGAACAATCAAGAATTATATTCTGTCATCTTCTCGAAGAACATTTGGCTGAGAAAT 271

194	QY	GAAGCTGATGTATGAATAATCAACGCGCTTGGAATPACGTTATCACTTTGAAACGGAGATTGA	253
272	Db	GGATCTGTAGACATTTGTTGAGCGCTTGCCATGCTTATCATTTTGAATAAACAATAGA	331
254	QY	TCATGCATTTGCAATGATTTTATGAACATATGGT-----GATAACTGGAA	298
332	Db	TGATATGTTGGATCAATTTTCAAGCAGATCCTTAACCTTTGAGGCTCAGAGTACAATGA	391
299	QY	TGGTGACCGCTCTTTCOTATGGTTCGGTCTTATGCGAAAGCAAGGATATATGTTACATG	358
392	Db	TTTACAAACTTTTATCGGTTCAATTTTCGACTATTGAGACAACATGTTTACAATATCTCCOC	451
359	QY	TGATGTTTTCAATAACTATAAGACAATAATGGAGCGTTCAAGCAATCGTTAGCTTAATGA	418
452	Db	AAAACTTTTTATTAGATTCCAAAGATGCAAAAGCAAAATTTAAGAATCTCTTTTGTAAACA	511
419	QY	TGTTGAAGGTTTGCTTGAGTGTGACGAACCACTTCTATGAGGCTACCTTGGGAGATTAT	478
512	Db	CATCAAGGCTCTTTGAACTTATACGAGCGCTGCAATGTAAAGCTCATGGAGAGATAT	571
479	QY	ATTAGAAGATGCTCTTGTTTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTT	538
572	Db	TTTGGAAAGGACATGCTTTTCTTACTGCTCATCTTGAATCT-----GCAGCTCC	622
539	QY	TTCTACAAACCGGCTCTTTTACCGAAATACAAACGGCAGCTTAAGCAACCCCTTTGGAA	598
623	Db	ACATTTGAAGTCAACCTCTGAGTAAGCAAGTGCACATGCCCTTGAGCAATCTCTCCATAA	682
599	QY	AAGGTTGCCAAGATAGAGCGCGCAGTACATTCCTTTCTATCAACAACAGATTCTCA	658
683	Db	GAGCATCCAGAGTTGACACACGCTACTTCACTCTATCTACGAGAGGAGAGACAGAA	742
659	QY	TAAACAGACTTTACTTAAACTTGTCTAAGTTAGAGTTCAATTTGCTTCACTGATGCACAA	718
743	Db	GAATGATGTGTTGCTTCAATTTGCAAACTGGACTTCAACTTACTTCAAGTTGTCACAA	802
719	QY	GGAAGAGCTCAGCCATGTTGCAAAATGGTGGAAAGCTTTGATATCAAGAAACGACCC	778
803	Db	ACAAGAACCTTAGTGAAGATACAAAGTGGTGGAAAGATTGGATTTGTGACAACACTTCC	862
779	QY	TTGTTTAAAGATAGAAATGTTGAATGCTACTTTTGGGACTTAGGTTAGGCTATCAGCC	838
863	Db	ATATGCTAGGATAGAGCAGTGGATGCTACTTTTGGACATGGGGGTATGCTGAACC	922
839	QY	ACAGTATTCGGGGCTAGAGTTTTTCTTCAAAAAGCTGTGTGTTATTAACCTCTTATAGA	898
923	Db	TCAATACTCTCAGGCTCGTGTATGCTTGAAGCTATAGCAATAGCAATTTCTTATAGTAGA	982
899	QY	TCACACTTATGATGGTATGTTAGTATCAAGAACTTAAGACTTTTCTGAGAGCTTTGA	958
983	Db	TGACACATTCGATGCTTATGGCATTTGCAAGAACTTGAGATCTACACCGATGCCATACA	1042
959	QY	AAGGTGGTCAATTATCATGTTCTAGACACACTTCACGAATACATGAACCCGATATACAAAT	1018
1043	Db	GAGGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATTACATGAATACTCAGTTTACAAGC	1102
1019	QY	ATTTCATGGATACATACAGAAATGGAGAAATTTCTTGCAAAAGGAGGGAAGAACATGCT	1078
1103	Db	ACTTTTATAGTCTCTACAATGATTTATGAATGGAGTTGTCCAAAGGATGGTAGATCTGTAT	1162
1079	QY	ATTTAACTCGGCAAGAAATTTGTGAACAGATTTGTTAGAAACCTCATGGTTGGAAGCAAA	1138
1163	Db	TGTTCACTACGGGAAGAAGAAATGAAGAAATCGTGAGAAACTATTTTGTGGAGACAAA	1222
1139	QY	ATTTGGGCAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTTGTATCATTTAC	1198
1223	Db	ATGGTTTCAATTGAAGGATATATGCGGCAGCTCTCTGAGTATCTTAGCAATGCATTAGCTAC	1282
1199	QY	TGGCGGTGCTAACCTGCTTACAAACAATTGTTATCTTGGCATGAGTATATATTCACAAA	1258
1283	Db	CAGCACTTATTACTGCTTACCACTACATCTTATTTGGGCATGAAGTCTGCTTAAACAAGA	1342

QY	1259	AGAGTCTGTGCAATGGCGTGTCTCTGCACCTCTCTTTTTTAGATATCTCAGGTATACTTGG	1318
DB	1343	AGATTTTG---AATGGTTGGCCAAGAACCCCTAAAAATCTTGAGGCTAAATGTGACGTTATG	1399
QY	1319	TCGACGCCCTTAATGATCTCATGCCACCAAGCGCCGACGAAGAAGAAAAACATAGTTTCATC	1378
DB	1400	CCGAGTCTATAGATGACATAGCCACCTTATGAGGTTTGAGAGGGTAGAGGTGAGTTGCCAC	1459
QY	1379	GAGCCCTTGAAAGTTTATATGAAGGAATATTAATGTCAAATGAGGAGTATGCCCAAACTTTGAT	1438
DB	1460	TGGAATTGAATGTTACATGAGAGATTAATGTGTGTATCCACAGAAAAAGCCATGGAAAAATT	1519
QY	1439	TTACAAGGAAGTAGAAGATGTGTGMAAGATATAAACCAGAGAGTACTCAACAACAAAAA	1498
DB	1520	CCAAGAAATGGCTGAGACAGACATGGAAGATGTAAATGAAGGAATCCTTCGACCAACTCC	1579
QY	1499	CATTCCAAAGCGCGTTATGATGCGTGTGATCTATTGTCGCCAGTTTCTTGAAGTTCAATA	1558
DB	1580	CGTCTCTACAGAGATTCTCACTCGCAATTCCTCAATCTTGCTCGCATTAATCGATGTTACTTA	1639
QY	1559	TGCAGGAA---AGGATAACTTCACACGATGGGAGACGAATCAAAACATCTCTAAAGTC	1615
DB	1640	TAAGCACAAATCAAGATGGATACACTCATCCGGAAGAAAGTACTAAACCTCATATTATTC	1699
QY	1616	TCTACTCGTTTATCCTATGAGTATATGA	1643
DB	1700	GTTGTTGGTGGACTCTATTGAAATTTAA	1727
RESULT 12			
AA	AAA38913	standard; DNA; 1644 BP.	
XX	AA	AAA38913;	
AC	AA	AAA38913;	
XX	AA	AAA38913;	
DT	25-AUG-2000	(first entry)	
XX	XX		
DE	XX		
XX	XX		
KW	XX	Tobacco 5-epi-aristolochene synthase Y527F DNA SEQ ID NO:7.	
KW	XX	Synthase; protein co-ordinate data; active site; modification; terpenoid;	
KW	XX	3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;	
KW	XX	isoprenoid; breeding programme; fragrance; flavour; pheromone;	
KW	XX	defensive agent; pigment; antitumour; steroid hormone;	
KW	XX	signal transduction pathway; bile acid; affinity purification;	
KW	XX	photoreceptor; enzymatic synthesis; nutrient supplement;	
KW	XX	immunological reagent; ds.	
XX	XX		
OS	XX	Nicotiana sp.	
OS	XX	Nicotinic.	
XX	XX		
PN	XX	WO200017327-A2.	
XX	XX		
PD	XX	30-MAR-2000.	
XX	XX		
PF	XX	17-SEP-1999; 99WO-US021419.	
XX	XX		
PR	XX	18-SEP-1998; 98US-C100993P.	
PR	XX	22-APR-1999; 99US-C130628P.	
PR	XX	23-AUG-1999; 99US-C150262P.	
XX	XX		
PA	XX	(KENT) UNIV KENTUCKY RES DEPT.	
PA	XX	(SALK) SALK INST BIOLOGICAL STUDIES.	
XX	XX		
PI	XX	Chappell J, Manna KR, Noel JP, Starks CM;	
XX	XX		
DR	XX	WPI; 2000-292839/25.	
DR	XX	P-PSDB; AAY90834.	
XX	XX		
PT	XX	Novel terpene synthase enzymes, useful for producing terpene	
PT	XX	hydrocarbons, e.g. fragrances or antitumor agents, are derived from known	
PT	XX	enzymes by specific amino acid alterations.	
PS	XX	Example 1; Page 349-352; 450pp; English.	

XX Novel terpene synthase enzymes, useful for producing terpene
PT hydrocarbons, e.g. fragrances or antitumor agents*, are derived from known
PT enzymes by specific amino acid alterations.
XX
PS Example 1; Page 349-352; 450pp; English.

XX The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 265-535 of a 548 amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha-carbon atoms (alphaC) that have interatomic distances, between each other, within tabulated ranges, have a centre point (within a sphere of radius 2.3 Angstrom) within tabulated ranges, and have an ordered arrangement of R groups (defining a side chain), excluding specific tabulated arrangements (tables given in the specification). (I), and related enzymes, are used to produce a wide range of terpenoids (e.g. cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances, flavours, pheromones, defensive agents, pigments, anticancer agents, components of signal transduction pathways, precursors of steroid hormones and bile acids, as photoreceptors and as co-factor side chains. Some synthases with little or no catalytic activity (and nucleic acids encoding them) are used as controls in the analysis of products formed by enzymatic synthesis; as nutrient supplements; for affinity purification of isoprenoids; or to develop immunological reagents or nucleic acids for monitoring expression of terpene synthase or inheritance of the gene in plant breeding programs. The new synthases may produce novel terpene products. AAY38910 to AAY38938 and AAY90831 to AAY90859 represent sequences used in the exemplification of the present invention

XX Sequence 1644 BP; 517 A; 322 C; 332 G; 473 T; 0 U; 0 Other;

Query Match 19.2%; Score 316.6; DB 3; Length 1644;
Best Local Similarity 51.9%; Pred. No. 1.1e-74;
Matches 846; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAAACCTATTGCGCCCATTCGCCAATTCCTCCAGCAATTTGGGAGATCAGTTT 74
DB 28 GAAGAAAGAGATTGTTGCGCCCGTGGCGGACTTCTCCCTAGTCTCTGGGGTATCAGTTC 87
QY 75 CTCATCTATCAAAAGCAAGTAGCAAGGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
DB 88 CTTTCATCTCCA---TTGATATCAGGTTGCGGAAAGTATGCTCAAGAGATTGAGCA 144
QY 135 GAAGTGGCGCACTACTAAAGAGAGCTTTGGATATTCCTATGAAACATGCCAATTTGGTTG 194
DB 145 TTGAAGGAACAAACGAGGAGTATGCTGTTAGCAACCGGAAGGAAATTTGGCCGATACATTG 204
QY 195 AAGCTGATTGATGAATTCACCGCTTGGAAATACCGTATCATCTTGAACGGGAGATTGAT 254
DB 205 AATTGATTGACATTATTGAAGCGCTTGGTATATCTCCACCTTTGAGAAGAAATTTGAT 264
QY 255 CATGATTGCAATGTATTATGAAACATATGTTGATAAAGTGGATGGTACCCCTCTTCC 314
DB 265 GAGATTTGGATCAGATTTACAAACCAAACTCAAATGCTGCAATGTTGTCACCTCTGCA 324
QY 315 TTATGGTTCGCTTATGCGAAGCAAGGATATTTATGATGATGTTTCAATAAC 374
DB 325 CTTCAATTTGATTTGCTCAGGCAACACCGGTTTCAACATCTCTCCTGAAATTTTCAGCAA 384
QY 375 TATAAGACAAAATGGAGCGTTTCAAGCAATGTTAGCTAATGATTTGAGGTTTGCCTT 434
DB 385 TTCCAGATGAAATTTGGCAAAATTCAGAGAGTCTCTTGGCTAGTATGCTTTAGGATATTA 444
QY 435 GAGTGTGTACGAACCACTTCTATGAGGGTACCTGGGGAGATTTATATAGAAGATGCTCTT 494
DB 445 AACTTGTATGAGCTTCACATGTAAGACTCATGCTGACGATATCTTAGAGAGCACTT 504
QY 495 GGTTTTACAGATCTGCTTTAGCATTTATGACAAAAGATGCTTTTCTCAAAACCCCGCT 554
DB 505 GCTTTTCTC-----CACTATCCATCTTGAATCTGCAGCTCCACATTTGAAATCTCCA 555
QY 555 CTTTTTACCGAATAACACGGGCATTAAGCAACCCCTTTGGAAAGGTTGCCAAGATA 614
DB 556 CTTAGGAGCAAGTACATGCCCTTTGAGCAATGTTGCAAGGGTGTTCCTAGATC 615
QY 615 GAGGGGGCGCAGTACATTCCTT---TCTATCAACAACAGATTTCTCATCAACAGACTTTA 671
DB 616 GAGACCCGATTTCTCATCTCATCAATCTATGACAAAGGAACATCGAAGATATATGTTA 675

RESULT 13
AAA38915
ID AAA38915 standard; DNA; 1644 BP.
XX

QY 672 CTTAAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCAATTCGCAACAAGAGAGCTCAGC 731
DB 676 CTTTCGATTTGCCAAATTTGGATTTCAACTTGTCTCCAGATGTTGCAACAACAGAACTTGTCT 735
QY 732 CATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGACGCACTTGTTTTAAGAGAT 791
DB 736 CAAGTATCAAGGTGGTGGAAAGATTTGGATTTGTAACAACACTTCCATATGCTAGAGAT 795
QY 792 AGAATTTTGAATCTACTTTTGGGGAGCTAGGTTTCAGGCTATGAGCCACAGATTTCCCGG 851
DB 796 CGAGTAGTTGAATGCTACTTTTGGGCATTAGGAGTTTATTTTGGCCTCAATACTCTCAA 855
QY 852 GCTAGAGTTTCTTCACAAAAGCTTGTGCTGTTTATTAACCTTATAGATGACACTTATGAT 911
DB 856 GCTCGCTGATGCTGCTGTTAAGACCATATCAATGATTTTCGATTTGTCATGACACCTTTGAT 915
QY 912 GCGTATGCTACTTATGAAGAACTTTAAGATCTTTACTGAAGCTGTTTGAAGGTTGGTCAAT 971
DB 916 GCTTACGGTACAGTTAAGAACTTTGAGGCATACACAGATGCCATACAAGATGGGATATC 975
QY 972 ACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA 1031
DB 976 AACGAAATTTGATCGGCTTCTGATTACATGAAATCACTTATAAAGCTATTCTAGATCTC 1035
QY 1032 TACACAAATGGAAGATTTCTTTGCAAGGAGGGAAGACAGATCTATTTAACCTCGGC 1091
DB 1036 TACAAGGATTTATGAAGAAGAAATTTCTAGTCCGGAAGATCTCATATTGCTGCCATGCA 1095
QY 1092 AAGAAATTTGTGAAGAGTTTGTTAGAAACCTGTGTTGAAGCAAAATTTGGGCAAAATGAG 1151
DB 1096 ATAGAAAGATGAAGAAGTAGTAGAAATTTATATGTCGAGTCAACATGGTTTATTGAA 1155
QY 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAAC 1211
DB 1156 GGATATATGCCACCTGTTCTGAATACCTAAGCAATGCACTAGCAACTACCATATTTAC 1215
QY 1212 CTGCTTACAAACCTTGTATCTTGGCATGATGATATTTTCAAAAAGAGCTCTCTCGAA 1271
DB 1216 TACCTCGCAACATCGTATTTGGGGCATGA---AGTCTGCCACGGAGCAAGATTTTGAG 1272
QY 1272 TGGGCTGTCTCTGCACCTCTCTTTTAGATACTCAGGTATCTTGGTTCGACGCCCTAAAT 1331
DB 1273 TGGTTGTCAAGATCCAAAATTTCTTGAAGCTAGTGTAAATTTATGTCGAGTTATCGAT 1332
QY 1332 GATCTCATGCCCAAGCGCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAAGT 1391
DB 1333 GACACGCCAGCTACGAGGTTTGAAAGACAGGGGACAAATTTGCAACTGGAATTTGAGTGC 1392
QY 1392 TATATGAGGATATATATGTCATGAGGAGTATGCCCAAACTTGTGATTTTACAGGAGTA 1451
DB 1393 TGATGAGAGATTTATGTTATATCAACAAAAGAGGCAATGGCTAAATTTCAAATATGGCT 1452
QY 1452 GAAGATCTGTGGAAGATATATAACCGAGAGTACCTCAACAACTTAAACAACTTCCAAAGCGC 1511
DB 1453 GAGACACATGGAAGAATATTTAAGAGGACTTCTTAGGCCACTCCCGCTCTTACAGAA 1512
QY 1512 TTATTTGATGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTTCAATATGACAGAAA---G 1568
DB 1513 TTTTAACTCTCTTCTCAATCTTCTGCTGCTGATTTGTTGAGGTTTACATATATACAACTTA 1572
QY 1569 GATTAATTTACACGATGTTGGGAGAGCAATACAAACATCTCATAAAGTCTTACTCTGTTAT 1628
DB 1573 GATGGATTTCTATCCGGAAGAGTCTTAAACCTCACTTATTAACCTTACTTGTGGAC 1632
QY 1629 CCTATGAGTAT 1639
DB 1633 TCCATCAAAAT 1643

AAA38915;
25-AUG-2000 (first entry)
Tobacco 5-epi-aristolochene synthase Y406X/L407X DNA SEQ ID NO:11.
Synthase; protein co-ordinate data; active site; modification; terpenoid;
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
isoprenoid; breeding programme; fragrance; flavour; pheromone;
defensive agent; pigment; antitumour; steroid hormone;
signal transduction pathway; bile acid; affinity purification;
photoreceptor; enzymatic synthesis; nutrient supplement;
immunological reagent; ds.
Nicotiana sp.
Synthetic.
WO200017327-A2.
30-MAR-2000.
17-SEP-1999; 99WO-US021419.
18-SEP-1998; 98US-0100993P.
22-APR-1999; 99US-0130628P.
23-AUG-1999; 99US-0150262P.
(KENT) UNIV KENTUCKY RES DEPT.
(SALK) SALK INST BIOLOGICAL STUDIES.
Chappell J, Manna KR, Noel JP, Starks CW;
WPI; 2000-292839/25.
P-PSDB; AAY90836.
Novel terpene synthase enzymes, useful for producing terpene
hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
enzymes by specific amino acid alterations.
Example 1; Page 358-360; 450pp; English.
The present invention describes an isolated terpene synthase (I)
comprising a region with at least 20% identity to region 265-535 of a 548
amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
-carbon atoms (alphaC) that have interatomic distances, between each
other, within tabulated ranges, have a centre point (within a sphere of
radius 2.3 Angstrom) within tabulated ranges, and have an ordered
arrangement of R groups (defining aa side chains), excluding specific
tabulated arrangements (tables given in the specification). (I), and
related enzymes, are used to produce a wide range of terpenoids (e.g.
cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
flavours, pheromones, defensive agents, pigments, antitumour agents,
components of signal transduction pathways, precursors of steroid
hormones and bile acids, as photoreceptors and as co-factor side chains.
Some synthases with little or no catalytic activity (and nucleic acids
encoding them) are used as controls in the analysis of products formed by
enzymatic synthesis; as nutrient supplements; for affinity purification
of isoprenoids; or to develop immunological reagents or nucleic acids for
monitoring expression of terpene synthase or inheritance of the gene in
plant breeding programs. The new synthases may produce novel terpene
products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
sequences used in the exemplification of the present invention
Sequence 1644 BP; 517 A; 319 C; 332 G; 470 T; 0 U; 6 Other;
Query Match 19.1%; Score 315.4; DB 3; Length 1644;
Best Local Similarity 51.7%; Pred. No. 2.4e-74;
Matches 843; Conservative 1; Mismatches 766; Indels 21; Gaps 5;
15 GAAGAAAACCTATTGCGCCCATTCGCAACTTTCCTCCCAAGCATTTGCGGAGATCAGTTT 74
26 GAAGAGAGATGTTGCGCCCGTCGCGGAGTTTCCCTAGTCTCTCGGGTGATCAGTTT 87

QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTAGTATGATTTAAAAAAA 134
DB 88 CTTTCATTTCTCCA---TTGATATCAGGTTGCGGAAAAGTATGCTCAAGAGATTGAAGCA 144
QY 135 GAAGTGGCGCAACTACTATAAAGAGCTTTGGATATTCCTATATGAAAACATGCCAATTGTTG 194
DB 145 TTGAAGGAACAAACACGAGGAGTATGCTTTAGCAACCGGAAGGAAATTTGGCGGATACATTG 204
QY 195 AAGCTGATTGATGAATTCACCGCTTGAATAACCGTATCATCTTGAACGGGAGATTGAT 254
DB 205 AATTTGATTGACATTATTGAACGCCCTTGGTATATCTCCACCTTTGAGAAGAAATTTGAT 264
QY 255 CATGCATTGCAATGATTATTGAACAATATGGTGAATACTGGAATGGTGACCGCTCTTCC 314
DB 265 GAGATTTTGGATCAGATTTTACAACCAAACTCAAACTGCAATGATTTGTGCACCTCTGCA 324
QY 315 TTATGGTTCCGCTTTATGGAAGAGAGATATTTATGTTACATGATGATGTTTTCATAAATAC 374
DB 325 CTTCAATTTGATGCTCAGGCAACACCGGTTTCAACATCTCTCCTCGAATTTTTCAGCAAA 384
QY 375 TATAAAGACAAAATGGAGCGTTTCAAGCAATGCTTAGCTAATGATGTTGAAGGTTTGTCTT 434
DB 385 TTCCAGATGAATGGAATGGCAATTCAGGAGTCTCTTGTAGTATGATGTTTAGGATTAATTA 444
QY 435 GAGTTGTAGAGAACAACTTCTATGAGGGTACCTGGGGAGATTATATTAGAAGATGCTCTT 494
DB 445 AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTTAGAAGACGCACTT 504
QY 495 GGTTTTACAGNATCTCGCTTTAGCATTATGACAAAAGATGCTTTTCTACAAACCCCGCT 554
DB 505 GCTTTTCTC-----CACTATCCATCTTGAATCTGCAGCTCCACATTTGAAATCTTCCA 555
QY 555 CTTTTCACGAAATCAACGGGCACCTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAATA 614
DB 556 CTTAGGAGCAAGTACACATGCTTTCAGCAATGTTTGCACAAAGGTTGTTCTTAGAGTC 615
QY 615 GAGCGGCGCAGTACATTCCTT---TCTATCAACAAAGATTTCTCATACAAAGACTTTA 671
DB 616 GAGACCCGATTTCTCATCTCATCAATCTATGACAAAGGAACAATCGAAGAAATATGTGTA 675
QY 672 CTTAAACTTGCTAAGTTAGATTCAATTTGCTTCACTCATTTGCACAAAGGAGAGCTCAGC 731
DB 676 CTTTCGATTGCGAAATTTGGATTTCACCTTGTCCAGATGTTGCACAAACAAAGAACTTGT 735
QY 732 CATGTGTGCAATGTGTGAAGCTTTCGATATCAAGAAACGCACTTGTGTTAAAGAGAT 791
DB 736 CAAGTATCAAGTGTGTGAAGATTGGATTTGTATCAACCACTTCATCATGCTAGAT 795
QY 792 AGAATTGTTGAATGCTACTTTTGGGACTPAGTTTCAGGCTATGAGCCACAGTATTCGCCG 851
DB 796 CGAGTAGTTGAATGCTACTTTTGGGCAATTAGGAGTTTATTTTGGAGCTCAATCTCTCAA 855
QY 852 GCTAGAGTTTCTTCAAAAGCTTGTGCTGTATTAACCTCTTATAGATGACATTTATGAT 911
DB 856 GCTCGGCTATGCTCGTCTTGAAGACCATATCAATGATTTTCGATGTCAGTACACCTTTGAT 915
QY 912 CGGTATGCTACTTATGAAGAACTTAAAGATCTTTTACGAGCTGTTGAAAAGTGGTCAATT 971
DB 916 CTTTACGCTACAGTTTAAAGAACTTGAGGCATACACAGATGCCATACAAAGATGGGATATC 975
QY 972 ACATGCTTTAGACACACTTCCAGAAATACATGAACCGGATATACAAATTTATTCATGGATACA 1031
DB 976 AACGAAATTTGATCGGCTTCTCTGATTACATGAAAATTCAGTTATAAAGCTATTTCTAGATCTC 1035
QY 1032 TACACAGAAATGGAAGATTTCTTCAAGAGGAGGAGAACAGATCTATTATTAATCGCGG 1091
DB 1036 TACAGAGGATTATGAAGAGAAATTTGTCTAGTGCAGGAGATCTCATATTGTCTGCCATGCA 1095
QY 1092 AAAGAAATTTGGAAGAGTTTGTGTAAGAACCTGATGTTGAAAGCAAAATGGGCAATGAG 1151
DB 1096 ATAGAAAGAAATGAAGAGATGATAGAAATTTAATATGTCGAGTCAACATGGTTTATTGAA 1155
QY 1152 GGACACATACCACCACTTGAAGAGATGATCCAGTTGTTAATCATTTACTGCGGCTGCTAAC 1211

Db 1156 GGATATATGCCACCTGTTCTCTGAATACCTAAGCAATGCACCTAGCACTACCAATATAC 1215
Qy 1212 CTGCTTACAAACACTGTTATCTTGGCATGAGTATATATTCACAAAGAGTCTGCGAA 1271
Db 1216 NNSNNSGCGCAACATCGTATATTTGGCATGA---AGTCTGCCACGAGCAAGATTTGAG 1272
Qy 1272 TGGGCTGTCTCTGCACCTCTCTCTTTTATAGATACTCAGGTATATCTTGGTGCAGCCCTAAAT 1331
Db 1273 TGGTGTCTCAAGAAATCCAAATCTTGAAGCTAGTGTAAATATATGTGAGTATTCGAT 1332
Qy 1332 GATCTATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCTATCGAGCCTTGAAGT 1391
Db 1333 GACACAGCCAGTACGAGGTTGAGAAAGAGCAGGGACAAATTCGAATCGAATTTGATGC 1392
Qy 1392 TATATGAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTGATTTACAGGAGTA 1451
Db 1393 TGCATGAGATTTATGTTATATCAACAAGGCAATAGGCTTAATTTCAAAATATGGCT 1452
Qy 1452 GAAGATGTGTGGAAGATATATAACCGAGAGTACCTCACAACCTTAAACCAATTCGAAGCCG 1511
Db 1453 GAGACAGCATGGAAGATATTAATGAAGGACTTCTTAGGCCACCTCCGCTCTACAGAA 1512
Qy 1512 TTATGTAGGCTGTGATCTATTTGGCAGTTTCTTGAAGTTCAATATGACAGAA---G 1568
Db 1513 TTTTAACTCTCTATTTCTCAATCTTCTCGTATGTTGAGGTTTACATATATACAACTTA 1572
Qy 1569 GATTAACCTTACACGATATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCTTTAT 1628
Db 1573 GATGATACATCTATCCGAGAAAGTCTTAAACCTCACAATATTAACTACTTGTGGAC 1632
Qy 1629 CCTATGAGTAT 1639
Db 1633 TCCATCAAAAT 1643

RESULT 14

AAA38910
ID AAA38910 standard; DNA; 1644 BP.
XX
AC AAA38910;
XX
DT 25-AUG-2000 (first entry)
DE Tobacco 5-epi-aristolochene synthase DNA SEQ ID NO:1.
XX
KW Synthase; protein co-ordinate data; active site; modification; terpenoid;
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW defensive agent; pigment; antitumour; steroid hormone;
KW signal transduction pathway; bile acid; affinity purification;
KW photoreceptor; enzymatic synthesis; nutrient supplement;
XX immunological reagent; ds.
OS Nicotiana sp.
XX
PN WO200017327-A2.
XX
PD 30-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US021419.
XX
PR 18-SEP-1998; 98US-0100993P.
PR 22-APR-1999; 99US-0130628P.
PR 23-AUG-1999; 99US-0150262P.
XX
PA (KENT) UNIV KENTUCKY RES DEPT.
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Chappell J, Manna KR, Noel JP, Starks CM;
XX
DR WPI; 2000-292839/25.
DR P-PSDB; AAY90831.

XX Novel terpene synthase enzymes, useful for producing terpene
PT hydrocarbons, e.g. fragrances or antitumour agents, are derived from known
PT enzymes by specific amino acid alterations.
XX
XX Disclosure; Page 336-339; 45Opp; English.
XX
CC The present invention describes an isolated terpene synthase (I)
CC comprising a region with at least 20% identity to region 265-535 of a 548
CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
CC -carbon atoms (aliphatic) that have interatomic distances, between each
CC other, within tabulated ranges, have a centre point (within a sphere of
CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered
CC arrangement of R groups (defining a side chain), excluding specific
CC tabulated arrangements (tables given in the specification). (I), and
CC related enzymes, are used to produce a wide range of terpenoids (e.g.
CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
CC flavour, pheromones, defensive agents, pigments, antitumour agents,
CC components of signal transduction pathways, precursors of steroid
CC hormones and bile acids, as photoreceptors and as co-factor side chains.
CC Some synthases with little or no catalytic activity (and nucleic acids
CC encoding them) are used as controls in the analysis of products formed by
CC enzymatic synthesis; as nutrient supplements; for affinity purification
CC of isoprenoids; or to develop immunological reagents or nucleic acids for
CC monitoring expression of terpene synthase or inheritance of the gene in
CC plant breeding programs. The new synthases may produce novel terpene
CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1644 BP; 518 A; 322 C; 332 G; 472 T; 0 U; 0 Other;
Query Match 19.1%; Score 315; DB 3; Length 1644;
Best Local Similarity 51.8%; Pred. No. 3e-74;
Matches 845; Conservative 0; Mismatches 765; Indels 21; Gaps 5;

Qy 15 GAAGAAAAACCTATTCCGCCCAATTCGCAACTTTCTTCAAGCATTTGGGGAGATCAAGTTT 74
Db 28 GAAGAGAGATTTGTCGCCCGCTCGCGCACTTCTCCCTAGTCTCTGGGGTATCAGTTC 87
Qy 75 CTATCTATCAAGCAAGTAGAGCAAGGGTGGACAGATAGTGAATGATTTAAAAAA 134
Db 88 CTTTCACTTCCCA---TTGATATATCAGGTTGCGGAAAGATATGCTCAAGAGATTGAACA 144
Qy 135 GAAATGCGGCACTACTATAAAGAGCTTTGGATATTCTTATGAACATGCCAATTTGTTG 194
Db 145 TTGAGGACAAACAGGAGATGCTGTTAGCAACCGAAGAAATTTGCCGATACATTG 204
Qy 195 AAGCTGATGATGAATCAAGCCCTTGGAAATACCGTATCACTTTGAACGGGAGATTGAT 254
Db 205 AATTTGATGACATATTGAACGCTTGGTATATCTTACCACCTTGAGAAAGAAATTCAT 264
Qy 255 CATGATTCGAATGTTATGAAACATATGTTGATACTGGATGGAATGGTACCCTCTCC 314
Db 265 GAGATTTGGATCAGATTTTACAAACCAAACTCAAACTGCAATGATTTGTGCACCTCTGCA 324
Qy 315 TTATGGTTCCTGTTATGCGAAGCAAGGATATTTATGTTACATGATGTTTCAATAAC 374
Db 325 CTTCAATTTGATTTGCTCAGGCAACAGGTTTCAACATCTCTCTGAAATTTTCAAGCAA 384
Qy 375 TATAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGGTT 434
Db 385 TTCCAAAGATGAAATGGCAAAATTCAGGAGTCTCTTGTCTAGTATGTTCTTAGGATTTATTA 444
Qy 435 GAGTTGTAGAGCAACTTCTATGAGGTACCTGGGAGATTATATTAGAGATGCTCTT 494
Db 445 AACTTGTATGAAGCTTTCATATGATGAGGATCTATGCTGACGATATCTTTAGAGAGCGCAT 504
Qy 495 GCTTTTACAGGATCTGCTTATGACATTATGACAAAAAGATGCTTTTCTCAAAACCCCGCT 554
Db 505 GCTTTCTC-----CACTATCCATCTTGAATCTGCAGTCCACATTTGAAATCTCCA 555
Qy 555 CTTTTTACGAAATACACGGGCACTAAAGCAACCCCTTTGGAAAGAGTTGGCAAGATA 614

556 CTTAGGGAGCAAGTGACATGCGCTTGAGCAATGTTTGCCACAAGGCGTGTCTCTAGATC 615
615 GAGGGGCGCAGTACATTCCTT---TCTATCAACAACAAGATTCCTCATACAGACTTTA 671
616 GAGACCCGATTCCTCATCTCATCAATCTATGACAGGAACAATCGAAGATAATGTGTTA 675
672 CTTAACTTGTCAAGTGAAGTGAAGTTCCTTCACTGATTCGACCAAGGAAGAGCTCAGC 731
676 CTTGATTTGCCAAATGCAATGCAATTCCTTCACTGCTCCAGATGTTGCACAAACAAGAACCTTGT 735
732 CATGTGTGCAAAATGTTGGAAGAGCTTTCGATATCAAGAAGACCCACTTGTGTTAAAGAT 791
736 CAAGTATCAAGGTGTGGAAGAGATTTGGATTTTGTAAACAACACTTCCATATGCTAGAT 795
792 AGAATTTGTTGAATGCTACTTTTGGGAGCTAGTTCAGGCTATGAGCCACAGTATTCCTCGG 851
796 CGAGTAGTGAATGCTACTTTTGGGAGCTAGGATTTATTTGAGCTCAATCTCTCAA 855
852 GCTAGAGTTTCTTCCAAAAAGCTGTGCTGTATTAACCTTTATAGATGACACTTATGAT 911
856 GCTGCGTCATGCTCGTTAAGACCATATCAATGATTTGCGATGTCGATGACACTTTGAT 915
912 GCGTATGCTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTTGAAAGGTGGTCAAT 971
916 GCTTACGGTACAGTTAAAGAACTTGAGGCATACACAGATGCCATACAAAGATGGGATATC 975
972 ACATGCTTAGACACACTTCCAGATACATGAACCGATATACAAATTTATTCATGGATACA 1031
976 AACGAAATGATCGGCTTCTGATACATGAATTCAGTTTAAAGCTATCTTAGATCTC 1035
1032 TACACAGAAATGGAAGAAATTTCTTCCAAAGGAGGGAACACAGATCTATTTAACTCGGGC 1091
1036 TACAAGGATTAAGAAAGAAATTTGCTAGTCCGGAAGATCTCATATTTCTGCTGCCATGCA 1095
1092 AAAGAAATTTGGAAGAGTTTGTAGAAACCTGATGGTTGAAGCAAAATGGGCAAAATGAG 1151
1096 ATAGAAGAAATGAAGAAAGTATGAAGAAATTAATGTCGAGTCAACATGGTTTATTGAA 1155
1152 GGACATACCAACCACTGAAGAGATGATCCAGTTGTAATCAATATTCGCGGTGCTAAC 1211
1156 GGATATATGCCACCTGTTCTGAAATACCTAAGCAATGCACTAGCAATCCACATATTAC 1215
1212 CTGCTTACAAACAATTTGTTATCTTGGCATGAGTGATATATTCACAAAGAGTCTGCGAA 1271
1216 TACCTCGGCAACAATCGTATTTGGGCAATGA---AGTCTGCCAGGAGCAAGATTTGAG 1272
1272 TGGGCTGTCTGCACTCTCTCTTTTAGATACCTAGGTATCTAGGTATCTAGGTCAGCCCTAAAT 1331
1273 TGGTTGTCAAAAGAAATCCAAAATTTCTGAAGCTAGTGTAAATATATATGTCGAGTTATCGAT 1332
1332 GATCTCATGACCCACAGCGCAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGT 1391
1333 GACACAGCAGTACAGGTTGAGAAAGCAGGAGCAAAATTCGCAACTGGAAATGAGTGC 1392
1392 TATATGAAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAGGAAGTA 1451
1393 TGCATGAGAGATTATGGTATATCAACAAGAGGCAATGCTTAAATTTCAAATATGGCT 1452
1452 GAAGATGTGGAAGATATAACCGAGAGTACCTCAACATTAACAAATTCAGAGCCG 1511
1453 GAGACAGATGGAAGATATAATGAAGAGCTTCTTAGCCCACTCTCTCTCTACAGAA 1512
1512 TTATTGAGGCTGATCTATTGTCAGTTTCTTGAAGTTCAATATGAGGAA---G 1568
1513 TTTTAACTCTTATCTCAATCTTCTGCTGATTTGAGGTTACATATATACAACTA 1572
1569 GATACTTACACGATGAGGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTAT 1628
1573 GATGATACACTCATCCGAGAGAAAGTCTTAAACCTCAATATTAACCTACTTGTGGAC 1632
1629 CCTATGAGTAT 1639
1633 TCCATCAAAAT 1643

RESULT 15
AAA38912

ID AAA38912 standard; DNA; 1644 BP.
XX
AC AAA38912;
XX
DT 25-AUG-2000 (first entry)
XX
DE Tobacco 5-epi-aristolochene synthase Y520F DNA SEQ ID NO:5.
XX
KW Synchase; protein co-ordinate data; active site; modification; terpenoid;
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
isoprenoid; breeding programme; fragrance; flavour; pheromone;
defensive agent; pigment; antitumour; steroid hormone;
KW signal transduction pathway; bile acid; affinity purification;
KW photoreceptor; enzymatic synthesis; nutrient supplement;
immunological reagent; ds.
XX
OS Nicotiana sp.
OS Synthetic.
XX
XX WO200017327-A2.
XX
PD 30-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US021419.
XX
PR 18-SEP-1998; 98US-0100993P.
PR 22-APR-1999; 99US-0130628P.
PR 23-AUG-1999; 99US-0150262P.
XX
PA (KENT) UNIV KENTUCKY RES DEPT.
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Chappell J, Manna KR, Noel JP, Starks CM;
XX
XX WPI: 2000-292839/25.
XX P-PSDB; AAY90833.
XX
PT Novel terpene synthase enzymes, useful for producing terpene
hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
enzymes by specific amino acid alterations.
XX
PS Example 1; Page 345-347; 450pp; English.
XX
CC The present invention describes an isolated terpene synthase (I)
comprising a region with at least 20% identity to region 265-535 of a 548
amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
-carbon atoms (alphaC) that have interatomic distances, between each
other, within tabulated ranges, have a centre point (within a sphere of
radius 2.3 Angstrom) within tabulated ranges, and have an ordered
arrangement of R groups (defining aa side chains), excluding specific
tabulated arrangements (tables given in the specification). (I), and
related enzymes, are used to produce a wide range of terpenoids (e.g.
cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
flavours, pheromones, defensive agents, pigments, antitumour agents,
components of signal transduction pathways, precursors of steroid
hormones and bile acids, as photoreceptors and as co-factor side chains.
Some synthases with little or no catalytic activity (and nucleic acids
encoding them) are used as controls in the analysis of products formed by
enzymatic synthesis; as nutrient supplements; for affinity purification
of isoprenoids; or to develop immunological reagents or nucleic acids for
monitoring expression of terpene synthase or inheritance of the gene in
plant breeding programs. The new synthases may produce novel terpene
products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
sequences used in the exemplification of the present invention
XX
SQ Sequence 1644 BP; 517 A; 323 C; 332 G; 472 T; 0 U; 0 Other;

Query Match 18.9%; Score 311.8; DB 3; Length 1644;
Best Local Similarity 51.7%; Pred. No. 2.2e-73;

Matches 843; Conservative 0; Mismatches 767; Indels 21; Gaps 5;			
Qy	15	GAAGAAAACCTATTTCGCCCGCATTTGCCAATTTCTCCCAAGCAATTTGGGAGATCAGTTT	74
Db	28	GAAGAAAGAAATTTGTCGCCCGCATTTGCCAATTTCTCCCAAGCAATTTGGGAGATCAGTTT	87
Qy	75	CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGAATGATTTTAAAAA	134
Db	88	CTTTCAATTTCTCA---TTGATTAATCAGGTTGCGGAAAGATGCTCAAGAGATTGAACA	144
Qy	135	GAAGTGGCGGAACCTACTAAAGAAAGCTTTGGATATTTCTTATGAACAATGCCAATTTGTTG	194
Db	145	TTGAAGGAACAAACGAGGAGTATGCTTTAGCAACCGGAAGGAAATTTGGCCGATACATTG	204
Qy	195	AAGCTGATTGATGAATCAACGCTTGAATACGATATCACTTCACTTGAACCGGAGATTGAT	254
Db	205	AATTTGATTGACATTTGAACGCTTTGGTATATCTTACCACTTTGAGAAAGAAATTTGAT	264
Qy	255	CATGCAATTCGAATGATTTATGAACAATATGTTGATAAATCTGGAATGGTGACCCCTCTTCC	314
Db	265	GAGATTTTGGATCAGATTTTACAACCAAACTCAAACCTGCAATGTTTGTCCACCTCTGCA	324
Qy	315	TTATGGTTCCGTTTATGCGAAGCAAGGATATTAATGTTACATGATGATGTTTTCATAAATC	374
Db	325	CTTCAATTTGATTTGCTCAGGCAACACCGGTTTCAACATCTCTCCTGAAATTTTTCAGCAA	384
Qy	375	TATTAAGCAAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTTGAAGGTTGCTT	434
Db	385	TTCCAAGATGAATGGCAATTCAGGAGTCTTTGCTAGTATGATCTTTAGGATTTATTA	444
Qy	435	GAGTTGTACGAACCACTTCTATGAGGTACCTGGGGAGATATATATAGAAGATGCTCTT	494
Db	445	AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTTAGAGACGACCT	504
Qy	495	GGTTTACAGATCTGCTTACGATTTATGAACAAAGATGCTTTTCTCAAAACCCCGCT	554
Db	505	GCTTTCTC-----CACTATCCATCTTGAATCTGCAGCTCCACATTTGAAATCTCCA	555
Qy	555	CTTTTACCAGAAATCAACGGGCATTAAGCAACCCCTTTTGGAAAGAGTTGCCAAGATA	614
Db	556	CTTAGGAGCAAGTACACATGCCCTTTAGGCAATGTTTGCACAAAGGTTGTTCTTAGATC	615
Qy	615	GAGCGGCGCAGTACATTCCTT---TCTATCAACAACAAGATTTCTCATACAAGACTTTA	671
Db	616	GAGACCCGATTTCTCATCTCATCAATCTATGACAGGAACAATCGAAGAAATATGTTGTTA	675
Qy	672	CTTAACCTTGTAGTTAGATTTCAATTTGCTTCAGTCAATTCGCAAGGAAGAGCTCAGC	731
Db	676	CTTCGATTTGCCAAATGGATTTCAACTTGTCTCCAGATGTTGCACAAACAAGAACTTGCT	735
Qy	732	CATGTTGCAATGCTGGAAGCTTTTCGATATCAAGAAACGACGACCTTGTTTTAAAGAT	791
Db	736	CAAGTATCAAGTGGTGGAAAGATTTGAAATTTGTAACAACTTCCATATGCTAGAGAT	795
Qy	792	AGAAATTTGTAATGCTACTTTTGGGACTAGGTTCAAGGCTATGAGCCACAGTATTTCCCGG	851
Db	796	CGAGTAGTTGAATGCTACTTTTGGSCATTAGAGTTTATTTTGGCCTCAATACCTCAA	855
Qy	852	GCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTATTAACCTCTTATAGATGACACTTAAGAT	911
Db	856	GCTCGCGTCATGCTCGTTAAGAACCATATCAATGATTTTCGATTTGATGACACCTTTGAT	915
Qy	912	GCATGCTACTTATGAAGAACTTAAGATCTTTTCTGAGCTGTTGAAGGTTGTTCAAT	971
Db	916	GCTTACGGTACAGTTAAAGAACTTTAGGCAATACAGATGCCATCAAAAGATGGGATATC	975
Qy	972	ACATGCTTTAGACACATTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA	1031
Db	976	AACGAAATTTGATCGCTTCTCTGATTTACATGAATAATCAGTTTAAAGCTATTTCTAGATCTC	1035
Qy	1032	TACACAGAAATGGAAGAAATTTCTTCCAAAGGAGGGAAGAACAGATCTATTTAACTCGCGC	1091
Db	1036	TACAAGGATTTATGAAGAAATTTGCTAGTGGCGGAAGATCTCATATTTGCTGCCATGCA	1095

Search completed: June 7, 2004, 09:34:50
Job time : 744 secs

Qy	1092	AAAGAAATTTGTGAAGAGTTTGTGTAGAAACCTGATGTTGTGAAGCAAAATGGGCAATGAG	1151
Db	1096	ATAGAAAGAAATGAAGAAAGTAGTAGAAATTTATATGTCGAGTCAACATGGTTTATTGAA	1155
Qy	1152	GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAAC	1211
Db	1156	GGATATATGCCACCTGTTTCTGTAATCCTAAGCAATGCACCTAGCACTACACATATTAC	1215
Qy	1212	CTGCTTACACAACTTGTGTATCTTTGGCAGTAGATGATATTTCCAAAGAGTCTGTCTGAA	1271
Db	1216	TACCTCGCAACACATCGTATTTGGGCATGA---AGTCTGCCACGAGCAAGATTTTGAG	1272
Qy	1272	TGGGCTGCTCTGCACTCTCTCTTTTAGATACTCAGGTATACTTGGTTCGACGCTTAAAT	1331
Db	1273	TGGTGTCAAGAATCCAAATTTCTTGAGCTAGTGTAAATTTATGTCGAGTTATCGAT	1332
Qy	1332	GATCTCATGACCCCAAGCGCGGCAAGAAAGAAACAATAGTTTCATCGAGCCTTGAAAGT	1391
Db	1333	GACACGCCACGTACGAGGTTTGAGAAAGCAGGGGACAAATTGCAACTGGAATTTGAGTGC	1392
Qy	1392	TATATGAAGAAATATAATGTCAATGAGGAGTATGCCCAAACTTTGATTTTACAAGGAAGTA	1451
Db	1393	TGCATGAGAGATTATGGTATATCAACAAAGAGGCAATGGCTAAATTTCAAATATGGCT	1452
Qy	1452	GAAGATGTGTGGAAGATATAAACCGGAGAGTACCTCAACAATAAAACATTTCCAGGCCG	1511
Db	1453	GAGACGATGGAAGATATTAATGAAGGACTTCTTAGGCCCACTCCCGTCTTACAGAA	1512
Qy	1512	TTATTTGATGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTTCAATATGCAAGAAA--G	1568
Db	1513	TTTTTAACTCTTATCTCAATCTTGTCTGATTTGTGAGGTTTACATTCACACAATCTA	1572
Qy	1569	GATAACTTTCACAGTATGGGAGAGCAATACAAACATCTCAATAAGTCTCTACTCGTTTAT	1628
Db	1573	GATGGATACACTCATCCGGAAGAGTCTTAAACCTCACAATTATTAACCTACTTGTGGAC	1632
Qy	1629	CCTATGAGTAT	1639
Db	1633	TCCATCAAAAT	1643

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: June 7, 2004, 09:06:39 ; Search time 151 Seconds
(without alignments)
6060.358 Million cell updates/sec
Title: US-09-763-822A-13
Perfect score: 1649
Sequence: 1 ccattggcacttacagaagaa.....ctatgagtatgaggatccc 1649
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Maximum Match 0%
Listing first 45 summaries
Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	364	22.1	1879	4	US-09-601-091-1 Sequence 1, Appli
2	357.6	21.7	2024	4	US-09-601-091-3 Sequence 3, Appli
3	357.6	21.7	2024	4	US-09-398-395A-51 Sequence 51, Appl
4	357.6	21.7	2024	4	US-09-887-586A-51 Sequence 51, Appl
5	357.6	21.7	2024	4	US-09-895-752-51 Sequence 51, Appl
6	357.6	21.7	2024	4	US-09-903-012B-51 Sequence 51, Appl
7	357.6	21.7	2024	4	US-09-900-797-51 Sequence 51, Appl
8	330.4	20.0	1944	4	US-09-398-395A-31 Sequence 31, Appl
9	330.4	20.0	1944	4	US-09-887-586A-31 Sequence 31, Appl
10	330.4	20.0	1944	4	US-09-895-752-31 Sequence 31, Appl
11	330.4	20.0	1944	4	US-09-903-012B-31 Sequence 31, Appl
12	330.4	20.0	1944	4	US-09-900-797-31 Sequence 31, Appl
13	320.6	19.4	1671	4	US-09-398-395A-1 Sequence 1, Appli
14	320.6	19.4	1671	4	US-09-887-586A-1 Sequence 1, Appli
15	320.6	19.4	1671	4	US-09-895-752-1 Sequence 1, Appli
16	320.6	19.4	1671	4	US-09-903-012B-1 Sequence 1, Appli
17	320.6	19.4	1671	4	US-09-900-797-1 Sequence 1, Appli
18	316.6	19.2	1644	4	US-09-398-395A-7 Sequence 7, Appli
19	316.6	19.2	1644	4	US-09-887-586A-7 Sequence 7, Appli
20	316.6	19.2	1644	4	US-09-895-752-7 Sequence 7, Appli
21	316.6	19.2	1644	4	US-09-903-012B-7 Sequence 7, Appli
22	316.6	19.2	1644	4	US-09-900-797-7 Sequence 7, Appli
23	315.4	19.1	1644	4	US-09-398-395A-11 Sequence 11, Appl
24	315.4	19.1	1644	4	US-09-887-586A-11 Sequence 11, Appl
25	315.4	19.1	1644	4	US-09-895-752-11 Sequence 11, Appl
26	315.4	19.1	1644	4	US-09-903-012B-11 Sequence 11, Appl
27	315.4	19.1	1644	4	US-09-900-797-11 Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-601-091-1
; Sequence 1, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1685)
US-09-601-091-1

Query Match 22.1%; Score 364; DB 4; Length 1879;

Best Local Similarity 52.9%; Pred. No. 1.7e-97;

Matches 865; Conservative 0; Mismatches 745; Indels 24; Gaps 3;

QY	29	TCGCCCATTCGCAACTTTCCTCCAGCAATTTGGGAGATCAGTTTCTCATCTATCAAAA	88
DB	65	TCGCCCATTCGCTAAATTTCCACCATCTGTTGGGATATCAATTCCTTTCTTACTCA	124
QY	89	GCAAGTAGAGCAAGGGTGGACAGATAGTGAATTTAAAAAAGAGTGGCAACT	148
DB	125	TGAAATTAATCAAGAAAAAGTTGAAGTTGATGATCAAGAGACAATTAGAAAT	184
QY	149	ACTAAAAGAGCTTTGGATATTCCTATGAACATGCCAAATTTGTTGAAGCTGATTGATGA	208
DB	185	GCTGTGGAAACTTCGCAATAGCACTCAAAA-----GCTGTGTTGATAGACGC	235
QY	209	AATTCACGCTTGGATACCGTATCTCTTGAACGGAGATTGATCATGCTTCAATG	268
DB	236	GATGCAACGATTGGGAGTGGCTTATCTATCGATAATGAAATGAAACATCAATCAAAA	295
QY	269	TATTTATGAACAT-----ATGTTGATACTGGAATGCTGACCGCTCTTCTT	316
DB	296	CATTTTGTGCAATCGTCCAAACAGATGATTAAGCAACACCTTACGTTGCTCTCT	355
QY	317	ATGTTTCCGCTTATGCGAAAGCAAGGATATATGTATGATGATGTTTCAATACTA	376

Db 356 TCGTTTTCGACTGTGAGCAACAAGGCCATTACATGTCTTCAGATGTGTTCAAGCAATT 415
Qy 377 TAAAGACAAATAATGAGCGTTCAGCAATCGTTAGCTAATGATGTTGAAGGTTGCTTGA 436
Db 416 CACCACCAAGATGGAATTCAGGAACAACATTACTAATGATGTCCAAGGATATTAG 475
Qy 437 GTTGACGAAGCAACTTCTATCAGGGTACTCGGGAGATTATATTAGAAGATCCTCTTGG 496
Db 476 TTTGATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATTCTTGAAGACTCTTAC 535
Qy 497 TTTTACAGATCTGCTTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCCGCTCT 556
Db 536 ATTTACCACCACTCATCTCGAGTCTATTCTCTCCAACCTTGAGCAATAATAAATCTCT 595
Qy 557 TTTTACGAATACAAAGCGCTAAAGCAACCCCTTTGABAAAGGTTGCCAGATAGA 616
Db 596 TAAGTTGAAGTTGGTGAAGCTTAATCAGCCTATTGCAATGACTTTACCAAGGATGGG 655
Qy 617 GCGCGCGAGTACATCTCTTCTATCAACAAGAGATTCTCAACAAGACTTTTACTTAA 676
Db 656 AGCTAGAAATACATATCATTTACGAACAACATGATGACACCACTTTGCTTTTGA 715
Qy 677 ACTTGCTAAGTTAGATTCATTTGCTTCAGTCAITTCGCAAGAGAGCTCAGCCATGT 736
Db 716 ATTTGCTAAATTTGATTTTAAATGCTGCAAAAGTTTCCAAAGAGAGCTTAGTGATCT 775
Qy 737 GTGCAATGCTGGAAGCTTTCGATATCAAGAGAGCGACCTGTTTAAGCATAGAT 796
Db 776 TACAAGGTGGTGAAGATTTGGATTTGCAATAAATATCCATATGCAAGAGACAGTT 835
Qy 797 TGTGAATCTACTTTTGGGGACTAGTTTCAGGCTATGAGCCACAGTATTCGCGGCTAG 856
Db 836 GGTGAGTGTACTCTGATATTAGAGTGTATTTTGGCCAAATATAGTCTGCGAG 895
Qy 857 AGTTTCTTCAAAAAGCTGTGCTGTTTAACTCTTTATAGATGACACTTTATGATGCTTA 916
Db 896 AAAAATGATGACAAAAGTACTCAACCTGACCTCCATTATTGAGGACACTTTTGTATGCTTA 955
Qy 917 TGGTACTTATGAGACTTTAAGTCTTTACTGAGCTGTTGAGGTGGTCAATTACATG 976
Db 956 TGCAACCTTTGACGAACCTTGTGACTTTCAATGATCAATCCAGAGATGGGATGCTAATGC 1015
Qy 977 CTTAGACACTTCCAGATACATGAAACCGATATACAAATTTTCATGATATACATACAC 1036
Db 1016 AATTGATTCAAATACACCATATATAGACCTGCTTATCAAGCTCTTCTAGCATTTACAG 1075
Qy 1037 AGAAATGGAAGATTTCTTGCAAAAGGAGGAAGAAAGATCTATTTAACTTGCAGCAAGA 1096
Db 1076 TCAAAATGGAACAAGTGTGTCAAAAGAGTAACTGGACCGGTATATCTATGCAAAAAA 1135
Qy 1097 ATTTGTGAAGAGTTTGTGTAAGAACCTGATGTTGAGCAAAATGCGCAATGA---GGG 1153
Db 1136 TGAGATGAAGAAAGTTGGTGAGAGCCCTATTTAAGGAACCCCAATGGTTGAATGTGA 1195
Qy 1154 ACACATACCACCACTGGAAGAGCATATCCAGTTGTAATCATTTACTGGCGGTGTACCT 1213
Db 1196 CCATATTCGAATAATGAGGAACAAGTGGAGATGCAATCGTAAGTCTGCTATATGAT 1255
Qy 1214 GTTTACAAACCTTGTTAICTTGGCATGAGTGATATATTACAAAAGAGTCTGTGCAATG 1273
Db 1256 GATATCAACAACTTGTCTGGTGTGATAGAAGAAATTTATATCCACGAGACTTTTGAATG 1315
Qy 1274 GGTGTCTCTGACCTCTCTTTTATAGATACTCAGGTATATCTTGGTGTGCGCCCTAATGA 1333
Db 1316 GTTGATGAATGATGTGTGATTTGTCAGCTTCCGCAATGATTGCCAGAGCAATGAAACGA 1375
Qy 1334 TCTCATGACCCCAAGGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCCCTTGAAGTTA 1393
Db 1376 TATTGTTGGACATGAAGATGAAACAGAAAGAGGACATGTAGCTTCACTTATTGAATGTTA 1435
Qy 1394 TATGAAGGAATAATAATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAGGAAGTAGA 1453

Db 1436 CATGAAGATTATGGAGCTTCAAAAGCAAGAGACTTACATTAAAGTTCTCTGAAAGAGGTAC 1495
Qy 1454 AGATGTGTGAAGATATAACCGAGAGTACCTCACAACTAAAAACATTCCAAAGCGGTT 1513
Db 1496 CAATGATGAAGAGACATAAACAACAATTTCTCGTCAACTGAAGTACCAATGTTTGT 1555
Qy 1514 ATTGATGGCTGTGATCTATTGTCGCCAGTTTCTTTGAAGTTCAATATGAGGAAAGATAA 1573
Db 1556 CCTTGAACGAGTCTTAAATTTGACACGTTGGCTGACACGTTATATAAAGAGAAAGATAC 1615
Qy 1574 CTTTACAGCTATGGGAGACGATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1616 ATATACAAACGCCAAAGGAAACCTTAAAAACATGATTAAATCAATFACATTAATGAAATCTGT 1675
Qy 1634 GAGTATATGAGGAT 1647
Db 1676 CAAAATATAATAT 1689

RESULT 2
US-09-601-091-3
; Sequence 3, Application US/09601091
; Patent No. 8342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)..(1678)
US-09-601-091-3

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

Qy 29 TCGCCCATTCGCCAACTTCTCTCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 89
Db 58 TCGCCCTTGGCTAATTTTCCACCATCTGTTGGGATATCATTTCTTCTTACTCA 117
Qy 89 GCAAGTAGACAGGGGTGGACAGATAGTGAATGATTAAAAAAGAGTGGCGCACT 148
Db 118 TGAATTTACTAATCAAGAAAAAGTTGAAGTTGATGATACAAAGAGACAATTAGAAAAAT 177
Qy 149 ACTAAAGAAGCTTTGGATATTCTTATGAAACATGCCAATTTGTTGAAGCTGATTGATGA 208
Db 178 GCTGGTGAACCTTGGCAATAGCACTCAAAA-----GCTTGTGTGATAGACGC 228
Qy 209 AATTCAAGCCCTTGAATACCGTATACCTTTGAACGGGAGATGATCATGCAATGCAATG 268
Db 229 GATGCAACGATTTGGAGTGGCTTATCATTTGATTAATGAATTAATGAACATCCATTCAAA 288
Qy 269 TATTATGAACAT-----ATGGTGAATTAATGGAATGGTGGCGCTCTTCCTT 316
Db 289 CATTTTGTATGATCGTCCAAACAGAAATGATTAATGACAAACCTTTAGCTGTGTCTCT 348
Qy 317 ATGTTTCCGTCTTATGCGAAAGCAAGGATATTATGTTACATGTGATGTTTTCATAACTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAGGCCATTACATGCTTCAGATGTGTTCAGCAATT 408
Qy 377 TAAAGACAAAAATGGAGCGGTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGCITGA 436

Db 409 CACCAACCAAGATGGGAATTC CAAGGAACACCTTACTAATGATGTCCAAGGATTATTGAG 468
QY 437 GTTGTACCAAGCAACTTCTATGAGGATCTCTGGGAGATTATATTAGAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATCTTTGAAGAAGCTCTTAC 528
QY 497 TTTTACACGATCTCGTCTTAGCATATGACAAAAGATGCTTTTCTACAAAACCCGCTCT 556
Db 529 ATTTACCACCACTCATCTCGAGTCTATTGTCTCCAACTTGAGCAATAATAAATCTCTCT 588
QY 557 TTTTACCGAATACAAAGGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAAATGA 616
Db 589 TAAGGTGAAGTGGTGAAGCCCTTAACCTCAGGCTATTTCGCATGACTTTACCAAGGATGG 648
QY 617 GCGGCGCAGTACATCTCTTTCTATCAACAAAGATTTCTCATACAAAGCTTTTACTTAA 676
Db 649 AGCTAGAAAATACATPACATTTAGAAAACAAATGATGACACCACTATTGCTTTTGA 708
QY 677 ACTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCAITGCAAGAGAGAGCTCAGCCATGT 736
Db 709 ATTTGCTAAATGGATTATAACATGCTGCAAAAGTTTCCAAAGAGAGCTTAGTATCT 768
QY 737 GTGCAATGGTGAAGCTTTTCGATATCAAGAAAGCAACCTTTGTTTGAAGATAGAAAT 796
Db 769 TACAAGGTGGTGAAGAGATTGGATTTTGCAAAATAAATATCCATATGCAAGAGACAGGTT 828
QY 797 TGTGTAATGCTACTTTTCGGGACTAGGTTTCAGGCTATGAGCCACATATTCGCGGCTAG 856
Db 829 GGTGAGTGTACTTCTCGATATTAGGAGTGATTTTGGAGCCAAAATATAGTCGTCGAG 888
QY 857 AGTTTCTTCAAAAGCTGTGCTTTATATCTTTATAGATGACACTATATGATCGGTA 916
Db 889 AAAAATGATGCAAAAGTACTCAACTGACCTCCATTATTGACGACACTTTTGTATGCTTA 948
QY 917 TGTGACTTATGAAGAACTTATAGATCTTTACTGAAAGCTGTGAAAGGTGTCATTAATG 976
Db 949 TGCAACCTTTGACGAACCTGTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
QY 977 CTTAGACACTTCCAGAAATACATAAACCAGATATACAAATTTATCATGATACATACAC 1036
Db 1009 AATTGATTCAATACAAACATATATAGAGACCTGCTTTATCAAGCTCTCTTAGACATTTACAG 1068
QY 1037 AGAAATGGAAGAATTTCTGCAAGAGGAGGAACAGATCTATTAACTGGGCAAGA 1096
Db 1069 TGAATGGAACAAGTGTGTCNAAGAAGGTAACTGGACCGTGATATCTATGCAAAA 1128
QY 1097 ATTTGTGAAGAGTTTGTGTAAGAACTGATGTTGAAGCAAAATGGCAAAATGA ---GGG 1153
Db 1129 TGAGATGAAAAAGTTGTGAGAGCCCTATTTTAAAGGAAACCCCAATGGTTGAATGTTGA 1188
QY 1154 ACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAACT 1213
Db 1189 CCAATTCCAAAATATGAGGAACAGTGGAGAATGCAATCGTAAGTGTGCTGCTATATGAT 1248
QY 1214 GCTTACAACAACTGTTATCTTGGCATGAGTGATATATTACAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAATCTGTTGTCGGTATAGAGAATTTATATCCAGAGACTTTTGAATG 1308
QY 1274 GGCTGTCTCTGCACCTCTCTTTTATAGTACTCAGGTATATCTGTCGACGCTCAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTGTTTCGAGCTTCCGCAATTTGATTTGCCAGAGCAATGA 1368
QY 1334 TCTCATACCCACAGGCGGAGCAAGAAACATAGTTTCATGAGCTTTGAAAGTTA 1393
Db 1369 TATTGTTGGCATGAAGATGACCAAGAAAGGACATGATGCTTCACTTATTGAATGTTA 1428
QY 1394 TATGAAGGAATATAATGTCATAGGAGTATGCCCAACCTTGATTTACAAGGAAGTAGA 1453
Db 1429 CATGAAGATATTGGAGCTTCAAGCAAGAGACTTTACATTAAGTTCTCTGAAAGAGGTAC 1488
QY 1454 AGATGTGTGAAGATATAAACCAGAGTACCTCAACTAAAAACATTTCCAAGGCCGTT 1513

Db 1489 CAATGCATGGAAGACATAAACAAACAATTTCTCCGTCCAACTGAAGTACCACATGTTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTGTCGCCAGTTTCTTGAAGTTCAATATGAGGAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGCTGTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
QY 1574 CTTTCAACAGTATGGGAGCGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACGCCCAAGGAAACCTTAAACATGATTAATCCAATCTAATTAATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 3
US-09-398-395A-51
; Sequence 51, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopodium esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-398-395A-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCATTCGCAACTTTCTCTCCAGCATTTGGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCATTCGCTAAATTTTCCACCATCTGTGGGATATCATTTCTCTTACTCTCA 117
QY 89 GCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAAAGAGTGGCGCAACT 148
Db 118 TGAATTAATAATCAAGAAAAGTTGAGTGTAGTACAAAGAGACAAATAGAAAAT 177
QY 149 ACTAAAGAGCTTTGGATATTCCTATGAACATGCCAAATTTGTTGAAGCTGATTGATGA 208
Db 178 GCTGTGTAACCTTCGACAAATAGCACTCAAAA-----GCTGTGTTGATAGACGC 228
QY 209 AATTCAACGCTTGGAAATACCGTATCACTTTGAACGGAGATTGATCATGATTCGAATG 268
Db 229 GATCAACGATTTGGGAGTGGCTTATCATTTTCGATTAATGAAATTTGAAACATCCATTCAAAA 288
QY 269 TATTTATGAAACAT-----ATGTTGATTAATGGAATGTTGACCGCTCTTCCTT 316
Db 289 CATTTTGTATGATCTGTCCTCAACAGAAATGATATGACAAACCTTTACGTTGTGTCTCT 348
QY 317 ATGGTTCCGTTTATGCGAAAGCAAGGATATTATGTTACATGTGATGTTTCAATAACTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAAGGCCATTACATGCTTTCAGATGTGTTCAAGCAATT 408

QY 377 TAAAGACAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGGTTGA 436
 Db 409 CACCAACCAAGATGGAAATTCAGGAACACCTTACTAATGATGTTCCAAAGGATTATTGAG 468
 QY 437 GTTGTACGAAGCAACTTCTATGAGGCTACTTGGGAGATTATATTAGAAGATGCTTCTGG 496
 Db 469 TTTGTATGAAGCATCATCTGAGAGTGGTAAATGAGGAGATCTTGAAGAGCTTCTTAC 528
 QY 497 TTTTACACATCTCTGCTTATGATATGACAAAGATGCTTTTCTACAAACCCGCTCT 556
 Db 529 ATTACCAACCATCTCTGAGTCTTATGCTTCTTCCAACTTGGACAAATTAATTAATCTCT 588
 QY 557 TTTTACCGAAATACAAACGGGCACTAAGCAACCCCTTTGGAAAGGTTGCCAAGAAATAGA 616
 Db 589 TAAGTTTGAAGTTGGTGAAGCTTAACTCAGCTATTTCGATGACTTTTACCAAGGATGG 648
 QY 617 GGGGGCGAGTACATCTCTTCTATCAACAAAGATTTCTATTAACAGACTTTTACTTAA 676
 Db 649 AGCTAGAAATACATATCCATTTACGAAACCAATGATGACACCCCACTTTCTTTTGA 708
 QY 677 ACTTGTAAAGTTAGAGTTCAATTTGCTTCAAGTCAATGCAAGGAGAGCTCAGCCATGT 736
 Db 709 ATTGTCTAATTTGAATTTTAACTGTCGAAAGTTTACCAAGAGAGCTTAGTGAAT 768
 QY 737 GTGCAAAATGGTGAAGAGCTTTCGATATCAAGAGAACGACCTTGTTTAAGAGATAGAT 796
 Db 769 TACAAGGTGGTGAAGAGATTTGGAATTTTGCATTTTCAATTAATATCCATATGCAAGAGACAGGTT 828
 QY 797 TGTGAATGCTACTTTTGGGAGTCTAGCTTCAAGCTATGAGCCACAGTATTCCTCGGCTAG 856
 Db 829 GGTGAGTGTACTTCTGATATTAGAGGTGATTTTGAAGCCAAATATAGTCTGCGAG 888
 QY 857 AGTTTCTTCAAAAAGCTTGTGCTGTTTAACTCTTATAGATGACACTTATGATGCGTA 916
 Db 889 AAAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTTGAAGCACCTTTTGAATGCTTA 948
 QY 917 TGTACTTATGAGACTTTAAGATCTTACTGAGCTGTTGAAGGTTGCTCAATTTACATG 976
 Db 949 TGCACCTTTGAGAACTTGTGACTTTCAATGATGCAATCCAGATGCGAGATGCTAATGC 1008
 QY 977 CTTAGACACTTCCAGATATACATGAACCGATATACAAATTTATTCATGATACATACAC 1036
 Db 1009 AATTGATTCAATACACCAATATATGAGACCTGCTTATCAAGCTCTCTAGACATTTACAG 1068
 QY 1037 AGAATGGAAGATTTCTTGAAGGAGGAGAACAGATCTATTAACTGCGGCAAGA 1096
 Db 1069 TGAATGGAACAGTGTGTTCCAAAGAGGTAACTGGACCGGTATATCTATGCAAAAA 1128
 QY 1097 ATTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGA--GGG 1153
 Db 1129 TGAGATGAAAAGTTGTTGAGAGCTTATTTAAGGAACCCATGTTGATGATGTA 1188
 QY 1154 ACATACCAACCACTGAAGAGATGATCGAGTTGTAATCATTAATCGGCGGTGTAACCT 1213
 Db 1189 CCATATTCAAAATATGAGGAACAAGTGGAGATGCAATCGTAAGTCTGCTATATGAT 1248
 QY 1214 GCTTACAAACACTTGTATCTTGGCATGATGATATATTCACAAAAGAGCTGTCGAATG 1273
 Db 1249 GATATCAACACTTGTGTTGGTGGTATAGAGATTTATATCCACGAGACTTTTGAATG 1308
 QY 1274 GGCTGCTCTGACCTCTCTTTTATGATCTCAGGTATACCTTGTGCGAGCTTAAATGA 1333
 Db 1309 GTTGTGATGATGAGTCTGTGATTTCTCGAGCTTCCGCAATGATTGCGAGCAATGAACGA 1368
 QY 1334 TCTCATGACCCCAAGGCGGAGCAAGAAAGAAACATAGTTCTATCGAGCTTTGAAAGTTA 1393
 Db 1369 TATTGTTGACATGAAGATGAACAAAGAAAGAGACATGATGCTTACTTATTGATGTTA 1428
 QY 1394 TATGAAGGAATATAATGTCAATGAGGATATGCCAAACCTTGTATTTACAGGAAGTAGA 1453
 Db 1429 CATGAAGATTATGGAGCTTCAAGCAAGAGACTTATACATTAAGTTTCTGAAAGAGTCA 1488

QY 1454 AGATGCTGTGGAAAGTATATAAACCGAGAGTACCTCACAACATAAAACATTTCCAAGCGCT 1513
 Db 1489 CAATGCTATGGAAGGACATAAAACAAATTCCTCGTCCAACTGAAGTACCAATGTTTGT 1548
 QY 1514 ATTGATGCTGTGATCTATTTTGTCCAGTCTTCTTGAAGTCTCAATATGAGGAAGATGA 1573
 Db 1549 CTTTGAACAGTCTTAAATTTGACACGCTGTGGCTGACACGCTTATATAAGGAGAAAGATAC 1608
 QY 1574 CTTTACACGATGAGGAGAGATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
 Db 1609 ATATTCAACCCCAAGGAAACCTTAAACACATGATTAAATCAATTAATTTGAATCTGT 1668
 QY 1634 GAGTATATGAGGAT 1647
 Db 1669 CAAAATATAAATAT 1682

RESULT 4
 US-09-887-586A-51
 ; Sequence 51, Application US/09887586A
 ; Patent No. 6495354
 ; GENERAL INFORMATION:
 ; APPLICANT: Chappell, Joseph
 ; APPLICANT: No. 64953541, Joseph P.
 ; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASES
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/887,586A
 ; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 09/398,395
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; PRIOR FILING DATE: 1999-08-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: RastSeq for Windows Version 3.0
 ; SEQ ID NO 51
 ; LENGTH: 2024
 ; TYPE: DNA
 ; ORGANISM: Lycopodium esculentum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (32)...(1675)
 ; OTHER INFORMATION: VENT germacrene C synthase
 ; US-09-887-586A-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
 Best Local Similarity 52.7%; Pred. No. 1.4e-95;
 Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCCATTGCCAATTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
 Db 58 TCGCCCCATTGGCTAATTTTACCCTATGTTGGGATATCATTTCTTCTTACTCA 117
 QY 89 GCAAGTAGACGAAGGGTGGAAACAGATAGTGAATGATTTAAAAAAGAGTGGCGCAACT 148
 Db 118 TGAATTAATTAATCAAGAAAAAGTTGAAGTTGATGAGTACAAAGAGACAAATAGAAAAAT 177
 QY 149 ACTAAAGAAAGCTTTGGATATCTCTATGAACATGCCAATTTGTTGAGAGCTGATTGATGA 208
 Db 178 GCTGGTGAAGAACTTGGCAATAGCAGCTCAAAA-----GCTTGTGTTGATAGACGC 228
 QY 209 AATTCAACGCTTGGATATACCGTATCACTTTTGAAGGAGAGATTGATCATGCAATGCAATG 268
 Db 229 GATGCAACGATTTGGAGTGGCTTATCATTTTCGATAATGAAATTTGAACATCAATTCAAAA 288
 QY 269 TATTATGAACAT-----ATGGTGAATCTGGAATGGTGAACCGCTCTCTCT 316
 Db 289 CATTTTGTATGATCGTCCCAACAGAAATGATTAATGACAAACCTTTACGTTGTGTCTCT 348
 QY 317 ATGGTTCGTTCTATGCGAAAGCAAGGATATTATGTTACATGATGATGTTTCAATAACTA 376

Db 349 TCCTTTTCGACTGTGAGGCAACAGGCCATTACATGCTTCAGATGTTCTCAAGCAATT 408
QY 377 TAAAGACAAAATAGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTCTTGA 436
Db 409 CACCAACCAAGATGGGAAATTCAGAGAAACACTTACTTAATGATGTTCCAGGATTTAGAG 468
QY 437 GTTGTACGAAGCAACTTCTATGAGGGTACCTGGGAGATTAATATAGAAAGATCTTTGG 496
Db 469 TTTGTATGAAGCATACATCTGAGAGTGGTAAATGAGGAGATTTCTTGAAGAAGCTCTTAC 528
QY 497 TTTTACACGATCTGCTTACCATTTATGACAAAGATGCTTTTCTACAAACCCGCTCT 556
Db 529 ATTACACCACTCATCTCGAGTCTATTGTCTCCTCAACTTGGAGCAATAATACTCTCT 588
QY 557 TTTTACCGAAATACACGGGCACTTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAAATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCTTAACTCAGCTTATTCGCACTTACCAGGATGGG 648
QY 617 GCGGGCGAGTACATCTCTTTCTATCAACAAAGATTTCTCATCAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTACGAAAACAAATGATGACACACCACTTTTGTGCTTTGAA 708
QY 677 ACTTGCTAGTTAGTTCAATTTGCTTCACTGCTATTCGACAGGAGGCTCAGCCATGT 736
Db 709 ATTGCTAAATGGATTTTAACTGCTGCAAAAGTTTCAACAAAGAGAGCTTAGTGATCT 768
QY 737 GTGCAATGTGGAAGCTTTGATATCAAGAAAGCGACCTTGTTTAAGAGATAGAAAT 796
Db 769 TACAAGTGTGGAAGATTTGATTTTGCATAATATCCATATGCAAGAGACAGTT 828
QY 797 TGTGAATGCTACTTTTGGGAGCTAGGTTGAGGCTATGAGCCACAGTATTCGCCGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATTTAGGAGTGTATTTTGAAGCAAAATATATGCTGCGAG 888
QY 857 AGTTTCTTCAAAAAGCTTGTGCTTATTAACCTTATAGATGACACTTATGATGCTA 916
Db 889 AAAATGATGACAAAAGTACTCAACCTGACCTCATTATGACGACACTTTTGTGCTTA 948
QY 917 TGTACTTATGAAGACTTAAAGTCTTACTGAGCTGTTGAAAGGTGTTCAATTACATG 976
Db 949 TGCAACTTTGACGACTTTGACTTTCAATGATGCAATCCAGAGATGGGATGCTAATGC 1008
QY 977 CTTPAGACACTTCCGAATATACATGAACCAACCATATACAAATATTTATGATGATACAC 1036
Db 1009 AATTGATTCATCAACACCATATATGAGACCTGCTTATCAAGCTCTCTAGACATTTACAG 1068
QY 1037 AGAAATGGAAGATTTCTTCAAGAGGAGGAGAACAGATCTATTTAAGTCCGCGCAAGA 1096
Db 1069 TGAATGGAACAAGTGTGTTCCAAAGAAAGGTAACTGGACCGTGTATATGCAAAA 1128
QY 1097 ATTTGTGAAGAGTTTCTTAGAACTGATGTTGAAGCAAAATGGGCAAAATGA--GGG 1153
Db 1129 TGAGATGAAAAGTTGGTGAGAGCTTATTTAAGGAAACCCATGTTGAATGATTTGTA 1188
QY 1154 ACACATACCAACCACTGAAGAGATGATCCAGTTGTAATCATTAAGTGGCGTGAACCT 1213
Db 1189 CCATATTCAAAATATAGAGAAACAAGTGGAGATGCAATCGTAAGTGTGCTATATGAT 1248
QY 1214 GCTTACAACTCTGTTATCTTGGCATGATGATATATTTACAAAAGAGTCTGCGAATG 1273
Db 1249 GATATCAACAACTTGTGCTGCTATAGAAAGATTTATATCCACGAGACTTTTGAATG 1308
QY 1274 GCGTGTCTGCACTCTCTTTTATAGATCTCAGGTATATCTGCTGAGCCCTAAATGA 1333
Db 1309 GTTGTGAATGAGTCTGTGATTTGTCAGCTTCCGCAATGANTCCAGAGCAATGAACGA 1368
QY 1334 TCTCATGCCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCCTTGAAGATTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAGAAAGAGGACATGTAGCTTCACTTATTGATGTTA 1428
QY 1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCCCAACCTTGTATTTACAGGAGTGA 1453

Db 1429 CATGAAGATTTATGGAGCTTCAAGCAAGAGACTTACATTAAGTTCTGAAAGAGGTAC 1488
QY 1454 AGATGTGTGAAAGATATAAACCAGAGAGTACCTCACAACTAAAAACATTTCCAAGGCCGTT 1513
Db 1489 CAATGATGGAAGGACATTAACAAACAATTTCTCCCGTCCAAGTGAAGTACCAATGTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTGTCGAGTTTCTTTGAGGATTTCAATATGAGGAGAAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGTTGCGCTGACACGCTTATATTAAGGAGAAAGATAC 1608
QY 1574 CTTCACACGTTATGGGAGACGAATACAAACATCTCAATAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACCGCAAGGAAACTTAAAAACATGATTAATCCAATACTAATTGAATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 5
US-09-895-752-51
; Sequence 51, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VENT germacrene C synthase
US-09-895-752-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 14e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
QY 29 TCGCCCCATTCACAACTTTCCCTCCAGAGATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCCCTTGGCTAATTTTCACCCATCTGTTGGGATATCATTTCTTCTTATATCA 117
QY 89 GCAAGTAGAGCAAGGGTGGAAACAGATGATGAATGATTAATAAAAAAGAAAGTGGGCAACT 148
Db 118 TGAATTTACTAATCAAGAAAAAGTTGAAGTGTGATGATACAAAGAGACAATTAGAAAAAT 177
QY 149 ACTAAAAAGAGTTTGGATATTCCTATGAAACATGCCAATTTTGTGAGCTGATGATGA 208
Db 178 GCTGTGGAATCTCGCAATAGCAGCTCAAAA-----GCTTGTGTTGATAGACGC 228
QY 209 AATTCAACGCTTGGAAATACCGTATCATTGAAACGGGAGATGATCATGCATTGCAATG 268
Db 229 GATGCAACGATTTGGAGTGGCTTATCATTTCCGATANTGAAATTTGAACATCCATCAAAA 288
QY 269 TATTATGAAACAT-----ATGGTGATACTGGAATGGTGACCGCTCTTCTCTT 316

Db 289 CATTTTGTGATCGTCGTCACAAACAGAAATGATAATGACAAACCTTTACGTTGTGTCCT 348
Qy 317 ATGGTTCGGTCTTATGCGAAGCAGGATATTTATGTTACATGTGATGTTTTCATTAACHTA 376
Db 349 TCGTTTTCGACTTGGAGCAACAGGCGCAATACATGCTTTCAGATGTGTTCAAGCAAT 408
Qy 377 TAAAGACAAAATGAGCGTTCGAAACATCGTTAGCTAATGATGTTGAAGGTTTGCTTGA 436
Db 409 CACCAACCAAGATGGAATTCAGGAACACCTTACTAATGATGTTCCAGAGATTTAGAG 468
Qy 437 GTTGACGAAGCAATCTTATGAGGGTACCTGGGAGATTAATATAGAAAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATTCCTGAAGAGCTCTTAC 528
Qy 497 TTTTACAGATCTCGTCTAGCATATGACAAAGATGCTTTTCTACAAACCCGCTCT 556
Db 529 ATTTACCACTCTCTCGAGTCTATTGCTCTCCAACTTTGAGCAATAATAAATCTCTCT 588
Qy 557 TTTTACCAAAATACAAACGGGCTAAAGCAACCCCTTTTGAAAAAGTTGCCAAGATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCTTAACTCAGCTATTCCATGACTTTTACCAAGGATGG 648
Qy 617 GCGCGCGAGTACATCTCTTCTATCAACAAAGATTTCTATACAAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAAATACATATCCATTTACGAAAACAATGATGACACACCACCACTTTGCTTGA 708
Qy 677 ACTTGCTAAGTTAGATTTCAATTTCTGCTCAGTTCATGTCACAGGAGAGCTCAGCATCT 736
Db 709 ATTTGCTAAATGGATTTTAAATGCTGCTGAAAAAGTTTCCAAAAAGAGCTTAGTATCT 768
Qy 737 GTGCAAAATGGTGAAGCTTTTCTGATATCAAGAAAGACCGACCTTTGTTTAAAGATAGAA 796
Db 769 TACAGGTGGTGAAGATTTGGAATTTGCAATAATAATCCATATGCAAGAGACAGGTT 828
Qy 797 TGTGAAATGCTACTTTTGGGAGCTAGGTTACGGTATGAGCCACAGATATTCGCGGCTAG 856
Db 829 GGTTCAGTGTACTTCTGATATTAGGAGTGTATTTGAGCCAAAATATAGTCGTCGAG 888
Qy 857 AGTTTCTTCAAAAGCTTGTGCTATATCTCTTATAGATGACACTTATGATCGCTA 916
Db 889 AAAATGATGACAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTGCTTA 948
Qy 917 TGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTTGAAAGTGTGCTCAATTACATG 976
Db 949 TGCAACCTTTGACGAATTTGACCTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
Qy 977 CTTAGACACTTCCAGAAATCATGAACCGGATATACAAATTAATCATGGAATACATAC 1036
Db 1009 AATTGATTCATACCAACCATATATGAGACCTGCTTATCAAGCTCTTCTAGACATTTACAG 1068
Qy 1037 AGAATGGAAGATTTCTTGCAAGGAGGGAAGAACAGATCTATTAACTGCGGCAAGA 1096
Db 1069 TGAATGGAACAAAGTGTGTCGAAGAGGTAACTGGGACCGTGTATATCTATGCAAAAA 1128
Qy 1097 ATTTGTGAAGAGTGTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGA---GGG 1153
Db 1129 TGAGATGAAGAGTGTGTAGAGCCCTATTTAAGGAAACCAATGTTGAATGATGTGA 1188
Qy 1154 ACACATACCAACCATGAAGAGCATGATCCAGTGTGATATCATCTGCGGCGTGTAACT 1213
Db 1189 CCATATTCAAAATATGAGGAACAAAGTGGAGATGCAATCGTAAGTGTGCTGTATATGAT 1248
Qy 1214 GCTTACAACTCTGTTATCTTGGCATGATGATATATTCACAAAGAGTCTGTCCAAGT 1273
Db 1249 GATATCAACAACTGCTTGGTGGTATAGAAATTTATATCCAGGAGACTTTTGAATG 1308
Qy 1274 GGCTGTCTGCACTCTCTTTTATAGATCTCAGGTATATCTGTCGAGCGGCTAAATGA 1333
Db 1309 GTTGATGATGATGCTGTGATGTTCTGAGCTTCGCAATTTGATGCCAGAGCAATGAAGA 1368
Qy 1334 TCTCATGACCCCAAGGCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTA 1393

Db 1369 TATTGTTGGACATGAGATGAACAAAGAAAGAGGACATGTAGCTTCACTTATTGAATGTTA 1428
Qy 1394 TATGAGGATATAAATGCTCAATGAGGAGTATGCCCAACCTTGTGATTTACAAGGAAGTAGA 1453
Db 1429 CATGAAGATATTGAGAGCTTCAAGCAAGAGACTTACATTAAGTTCCTGAAAGAGGTAC 1488
Qy 1454 AGATGTGTGAAAGATATAAAACCGAGAGTACCTCACAACTAAAAACATTTCCAAAGCCGTT 1513
Db 1489 CAATGATGGAAGAGACATAAAACCAAACTCTCCGCTCCAACTGAAGTACCAATGTTTGT 1548
Qy 1514 ATTGATGGCTGTGATCTATTTTGTGCCAGTTCCTTGAAGTTCATATGAGGAAAGATAA 1573
Db 1549 CCTTGAACGAGTTCCTAAATTTGACACGCTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
Qy 1574 CTTTACACGATGAGGAGACGAATCAAAACATCTCATAAAGTCTCTACTCGTTTATCCCTAT 1633
Db 1609 ATATTCAACCGCAAGGAAAACTTAAAAACATGATTAATCCAATCTAATTAATGATCTGT 1668
Qy 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 6
US-09-903-012B-51
; Sequence 51, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65696561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; PRIORITY FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIORITY FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIORITY FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIORITY FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germaecene C synthase
US-09-903-012B-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

Qy 29 TCGCCCAATGCCAATCTTCTCCCAAGCATTTGGGAGATCAGTTCTCATCTATCAAA 88
Db 58 TCGCCCAATGCCAATCTTCCCAATCTGTTGGGATATCATTTCTTCTTACTCA 117
Qy 89 GCAAGTAGACAAAGGGGTGGAACAGATGTAATTAATAAAAGAAAGTCGGCACT 148
Db 118 TGAATTAATTAACAAGAAAAGTTGAAGTTGATGAGTACAAAGAGACAAATAGAAAAT 177
Qy 149 ACTAAAAGAGCTTTGGATATTCCTATGAACATGCCCAATTTGTTGAAGCTGATGAGA 208
Db 178 GCTGGTGAAGAACTGCGCAATAGCACTCAAAA-----GCTTGTGTTGATAGACG 228
Qy 209 AATTCACGCTTTGGATACCGTATCACTTTGAACGGGAGATTTGATCATGATTCGAATG 268
Db 229 GATGCAACGATGGGAGTGGCTTATCATTTTCGATAATGAAATGAAACATCAATCAAAA 288

QY 269 TATTATGAAACAT-----ATGGTGATAACTGGAATGGTGACCGCTCTTCCTT 316
Db 289 CATTTTGTGATGTCATCGTCCAAACAGAAATGATAATGACAAACACCTTTACGTTGTGCTCT 348
QY 317 ATGGTTCGCTTATGCGAAGCAGGATATTTATGTTACATGATGATGTTTTCNAATACTA 376
Db 349 TCGTTTTCGACTGTGAGGCAACNAGGCCATACATGCTTCCAGATGTTTTCNAGCAAT 408
QY 377 TAAAGACAAAATGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGCTTGA 436
Db 409 CACCAACCAAGATGGGAATTCAGGAAACACTTACTATGATGTTCCAGAGATATTGAG 468
QY 437 GTTGACGAACCACTTCTATGAGGTAACCTGGGAGATTAATATTAGAAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATCTTGAAGAGCTCTTAC 528
QY 497 TTTTACAGATCTGCTCTTAGCAATATGACAAAGATGCTTTTCTACAAAACCCGCTCT 556
Db 529 ATTTACCAACCACTCATCTCGAGTCTATTGTCTCCAACTTGGCAATTAATAAATCTCT 588
QY 557 TTTTACCAAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAGAGTTGCGCAAGATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCCTTAACTCAGCTATTCGATGACTTTTACCAAGGATGGG 648
QY 617 GCGGCGCAGTACATTCCTTTCTATCAACAAGATTTCTCATTAACAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTACGAAAACAATGATGACACACCACCAATTTGCTTTGAA 708
QY 677 ACTTGCTAAGTTAGATTCATTTGCTTCAGTCAITGCAACAGGAGAGCTCAGCCATGT 736
Db 709 ATTTGCTAAATGGATTTTAAATGCTGCAAAAGTTTCCAAAGAGAGCTTAGTGATCT 768
QY 737 GTGCAATGGTGAAGCTTTTCGATATCAAGAAAGACGACCTTTTAAAGATAGAAT 796
Db 769 TACAAGTTGGTGAAGATTTGATTTTCCAAATAAATATCATATGCAAGAGACAGTT 828
QY 797 TGTGAAATGCTACTTTTGGGAGTACTAGTTTCAAGCTATGAGCCACAGTATTCGCGGCTAG 856
Db 829 GGTGAGTGTACTCTCGATATTAGGAGTGTATTTTGAGCCAAATAATATAGTCTGCGAG 888
QY 857 AGTTTCTTCAAAAAGCTGTGTTGTTAATCTTATAGATGACACTTATGATGCGTA 916
Db 889 AAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATTGAGGACACTTTTGTATGCTTA 948
QY 917 TGGTACTTATGAAGAACTTTAAGATCTTTTACTGAAGCTTTGAAAGTGGTCAATTTACATG 976
Db 949 TGCAACCTTTGACGAACCTTGTGACTTTCAATGATGCAATCCAGAGATGGATCTAATGC 1008
QY 977 CTTAGACACACTTCCAGAAATACATAAACCAGATATACAAATTTATTCATGATACATACAC 1036
Db 1009 AATTGATTCAATACAAACCATATATGAGACCTGCTTATCAAGCTCTTCTAGACATTTACAG 1068
QY 1037 AGAAATGGAGAAATTTCTTGCAAGAGGAGAGAACAGATCTATTAACTGCGGCAAGA 1096
Db 1069 TGAAATGGAAACAAAGTGTGTGTCAAAGAGGTAAACTGGGACCGGTATATCTATGCAAAAA 1128
QY 1097 ATTTGTGAAAGATTTGTTAGAAACCTGTGTTGAAGCAAAATGGGCAATGCA---GGG 1153
Db 1129 TGAGATGAAAAGTTGGTGAGAGCTTATTTAAGGAAACCCAAATGGTGTGAATGTTGTA 1188
QY 1154 ACACATACCAACCACTGAAAGAGCATGATCCAGTGTGTAATCATTAATPACTGGGGTCTAACCT 1213
Db 1189 CCAATTTCCAAATATGAGGAACAAGTGGAGATGCAATCGTAAGTCTGGCTATATGAT 1248
QY 1214 GCTTCAACAACTTGTATCTTTGGGCAATGATGATATATTACAAAAGATCTCTCGAATG 1273
Db 1249 GATATCAACAACTTGTGTCGGTATAGAAATTTATATCCACAGACTTTTGAATG 1308
QY 1274 GCCTGCTCTGCACTCTCTTTTATAGATACTCAGGTATATCTTGGTCGACGCTAATGA 1333
Db 1309 GTTGAATGATGATCTGTGATTTGCTCGAGCTTCGCAATGATTTGCCAGAGCAATGAACGA 1368

QY 1334 TCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTA 1393
Db 1369 TATTGTTGGACATGAAAGATGAACAAAGAAAGAGGACATGATAGCTTCACTTATTGAATGTTA 1428
QY 1394 TATGAAGGAATATAATGTCATGAGGAGTATGCCAAACCTTGATTTACAAAGGAAGTAGA 1453
Db 1429 CATGAAGATATTGAGAGCTTCAAGCAAGAGACTTACATTAACTTCTCTCAAGAGAGTAC 1488
QY 1454 AGATGTGTGAAAGATATAAACCCGAGAGTACCTCACAACTAAAAACAATTCCAAGGCCGTT 1513
Db 1489 CAATGCAATGAAAGGACATATAACAAACAATTCCTCCGTCCAACTGAAGTACCAATGTTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTGTCGCCAGTTCTTGAAGTCAATATGCAGGAAGGATAA 1573
Db 1549 CCTTGAACGAGTTCTTAAATTTGACACGTGTGGCTGACAGCTTATATAAGGAGAAAGATAC 1608
QY 1574 CTTACACAGTATGGGACAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTATCTCTAT 1633
Db 1609 ATATTCAACCGCCCAAGGAAACTTAAAAACATGATTTATCCAACTACTTAATTGAATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 7
US-09-900-797-51
; Sequence 51, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 66457621, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VPNT germacrene C synthase
US-09-900-797-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
QY 29 TCGCCCATTCGCCAACTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCTTGGCTAATTTTCAACCATCTGTTTGGGATATCATTTCCITTTTATACTCA 117
QY 89 GCAGTAGACAAAGGGTGGACAGATAGTGAATTTAAAAAAGAGTGGCGCAACT 148
Db 118 TGAATTAATCTAATCGAAGAAAGTTGAAGTTGATGATCAAGAGACAAATTAGAAAAAT 177
QY 149 ACTAAAAAGAGCTTTGGATATTCCTATGAAAACATGCCAAATTTGTTGAAGCTGATGA 208
Db 178 GCTGGTGGAACTTTCGACATAGCACTCAAAA-----GCTTGTGTTGATAGACGC 228
QY 209 AATTCAACGCCCTCGAATACCGTATCACTTTGAACGGGAGATGATCATGCATTGCAATG 268

Db 229 GATGCAACGATTGGGAGTGGCTTATCATTTTCGATAATGAAATTTGAAACATCCATTCAAAA 288
Qy 269 TATTATTAACAAAT-----ATGGTGAATACTGGAATGGTACGGCTCTTCCTT 316
Db 289 CATTTTGTATGATCGATCGTCCAAACAGAAATGATGAAACAAACCTTTACGTTGTGTCTCT 348
Qy 317 ATGGTTTCGTCTTATGCGAAACCAAGGATATTATGTTATCATGTGATGTTTTCATTAACATA 376
Db 349 TCGTTTTCGACTTGTGAGCAACACAGGCCATTCATGCTTCAGATGTGTTCAAGCAAT 408
Qy 377 TAAAGACAAAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGTCTTGA 436
Db 409 CACCAACCAAGATGGGAAATTCAGGAAACACTTACTAATGATGTCCAAGGATTTATTGAG 468
Qy 437 GTGTGACAGCAACTTCTATCAGGGTACCTGGGAGATTATATAGAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGCCTAATGAGGAGATTCCTTGAAGAGCTCTTAC 528
Qy 497 TTTTACAGCATCTCGTCTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCCGCTCT 556
Db 529 ATTTACCACTCATCTCGAGTCTATTGCTCCAACTTGAGCAATTAATAAATCTCT 588
Qy 557 TTTTACCGAAATPACACGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGATAGA 616
Db 589 TAAGGTGAAGTTGGTGAAGCTTAACTCAGCCTATTGCGATATTCGATTTACCAAGGATGG 648
Qy 617 GCGCGCGAGTACATCTCTTCTATCAACAAAGATTTCTCATAAAGATTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCATTTACGAAAACAAATGATGCACACCACTTTGCTTTGAA 708
Qy 677 ACTTCTAGTTAGAGTTCAATTTGCTTCACTTCAAGCAAGGAGGCTCAGCCATGT 736
Db 709 ATTTCTAAATTTGGATTTTAAATGCTGCAAAAGTTTCCAAAGAGAGCTTAGTATCT 768
Qy 737 GTGCAATGCTGAAAGCTTTGATATCAAGAAAGACGACCTTTTGAAGATAGAT 786
Db 769 TACAAGGTGGTGAAGATTTGGATTTGCAAAATAATATCCATATGCAAGACACAGTT 828
Qy 797 TGTGAATGCTACTTTTGGGACTAGTTTCAGCTATGAGCCACAGATTTCCCGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATAGAGTGTATTTTGAGCCAAATATATAGTCTGTCGAG 888
Qy 857 AGTTTCTTCAAAAAGCTTGTCTGTATACTTATAGATGACACTTTATGATGCTTA 916
Db 889 AAAATGATGACAAAGTACTCACTGACCTCATTTATGACGACACTTTTGTATGCTTA 948
Qy 917 TGGTACTTATGAAGAACTTTAAGATCTTTTACTGAGCTTTTGAAGGTTGTCATTAATG 976
Db 949 TGCACCTTTGACGAACTTGTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
Qy 977 CTAGACACTTCCAGAAATACATGAAACCGATATACAAATTTATGATGATACATAC 1036
Db 1009 AATTGATTCATACAACTATATGAGCTGTTTATCAAGCTCTTCTAGACACTTTTACAG 1068
Qy 1037 AGAAATGGAAGAAATTTCTTCAAGAGGAGGAAAGACAGATCTTTTAACTGCGCGCAAGA 1096
Db 1069 TGAATGGAACAGTGTGTCGAAGAGAGTAACTGGACCGTGTATCTATGCAAAAA 1128
Qy 1097 ATTTGGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAAATGGGCAATGA---GGG 1153
Db 1129 TGAGATGAAAAAGTTGGTGAGAGCTTATTTAAGGAAACCAATGGTTGAATGATTGTA 1188
Qy 1154 ACACATACCAACCACTGAGAGCATCATCAGTTGTATCATCTACTGGCGGTGCTAACCT 1213
Db 1189 CCATATTTCCAAATATGAGAAACAGTGGAGATGCAATCGTAAGTCTGGCTATATGAT 1248
Qy 1214 GCTTACCAACAACTGTTATCTTGGCATGAGTATATTTTACAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACTTGTGTCGGTATAGAGAAATTTATATCCACAGAGCTTTTGAATG 1308
Qy 1274 GCGTGTCTCTGCAACCTCTCTTTTATAGATCTCAGGTATATCTGGTGCAGCGCTTAAATGA 1333

RESULT 8
US-09-398-395A-31
; Sequence 31, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09398,395A
; PRIORITY FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-398-395A-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Pred. No. 1.5e-87;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

Qy 14 AGAAGAAAAACCTATTCGCCCATTTGCCAATCTTCTCCAGCATTTTGGGAGATCAGTT 73
Db 92 AGAGAGAGAGATTGTTTCGCCCATCTAGTACTTCTCTCCAGCTTTGGGGTATCGTTT 151
Qy 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTCTTCTCCCTCGACAATCAGATTGCTGAAAAATATGCTCAAGAGATCGAACTTT 211
Qy 134 AGAAGTGGCGCACTACTAAAGAGCTTTGGATATTCTATGAACATGCGCAATTTGTT 193
Db 212 GAAGGAACAATCAAGAATTATATTATCTGATCTTCTCGAAGAACATTCGCTGAGAAAT 271

QY	194	GAAGCTGATTGATGAAATTTCAACGGCTTTGGAAATACCGTATACATTTTGAAACGGGAGATTGA	253
Db	272	GGATCTGATAGACATGTTGTAGCCCTTGGCAITTGCTTATCATTTTGAAATAACAATATGA	331
QY	254	TCATGCAATTCGAATGATTTTATGAAACAATATGGT-----GATAACTCGAA	298
Db	332	TGATATGTTGGATCAATTTTACAAGACAGATCCTAACTTTGAGGCTCAGAGATCAATGA	391
QY	299	TGGTGACGGCTTCTTCCTATATGGTTCCGTCTTATCGAAAGCAAGATATTTATGTTACATG	358
Db	392	TTTACAAACTTTTATCCGTTCAAATTCGACTATTGAGACAAATGGTTACAATATCTCCCC	451
QY	359	TGATGTTTTTCAATAAATCTATAAGACAAAATGGAGCGCTCAAGCAATCGTTAGCTAATGA	418
Db	452	AAAACTTTTATTTAGATTCOMAGATGCAAAAGGCAAAATTTAAGAAATCTCTTTGTAAACGA	511
QY	419	TGTTGAAGGTTTGCTTGAGTTGTACGAAGCAAACTTCTATGAGGTCACCTGGGAGATTTAT	478
Db	512	CATCAAGGGCTTTTGAACATTTATACGAAGCCTCGCATGTAAGGACTCATGGAGAGATAT	571
QY	479	ATTAGAAGATGCTCTTGCTTTTACAGATCTCGTCTTTAGCAATTTANGACAAAAGATGCTTT	538
Db	572	TTTGGAAAGAGGCACTTCTCTTTCTCTACTGTCTCATCTTGAATCT-----GCAGCTCC	622
QY	539	TTCTACAAAACCCGCTCTTTTACCGAAATACACGGGCACATAAGCAACCCCTTTGGAA	598
Db	623	ACATTTGAAGTCACTCTGAGTAGCAAGATGACACATGCCCTTGAGCAATCTCTCCATAA	682
QY	599	AAGGTTGCCAAGAATAGAGCGCGCGAGTACATTTCTCTTTCTATCAACAACAAGATTTCTCA	658
Db	683	GAGCAITTCGAAGAGTTGAGACACGCTACTTCACTCTCTATCTACGAAGAGGAGAACAGAA	742
QY	659	TAAACAAGACTTTTACTTAACTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCAATGCACAA	718
Db	743	GAATGATGTGTGTGCTTCAAATTTGCMAAATCGGACTTCAAATTAATTCAGATGTGCACAA	802
QY	719	GGAAAGCTCAGCAATGCTGTGCAAAACGGTGGAAAGCTTTCGATATCAAGAAGAACGCACC	778
Db	803	ACAAAGAACTTAGTGAAGTATCAAGTGGTGGAAAGATTTGGATTTTGTGACAACTTCC	862
QY	779	TTGTTTAAAGAGATAGAAATTTGTAATGCTACTTTTGGGGACTAGGTTTCAGGCTATGAGCC	838
Db	863	ATATGCTAGGATAGACAGTGGAAATGCTACTTTTGGACGNTGGGGGTGATGCTGNACC	922
QY	839	ACAGTATTCGGGGCTAGATTTTCTTCAAAAAGCTGTGTTGTTTATTAATCTTTATAGA	898
Db	923	TCAACTACTCTCAGGCTCGTCTCATGCTTGCTTAAGACTATAGCAATGATTTCTATAGTAGA	982
QY	899	TGACACTTATGATGCGTATGTAATGATGAAGAACTTAAATCTTTTACTGAAAGCTGTGA	958
Db	983	TGACACAAATCGATGCTTATGGCATGTGTCAAAGAACTTGAGATCTACACCGATGCCATACA	1042
QY	959	AAGTGGTCAATTTACATGCTTTAGACACACTTCCAGATATACATAAACCAGATATACAAAT	1018
Db	1043	GAGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATTTACATGAATCAGTTTACAAAGC	1102
QY	1019	ATTTCATGGATACATACACAGAAATGGAAAGAAATTTCTTGCAAAAGAGGGAAGAACAGATCT	1078
Db	1103	ACTTTTAGATCTCTACAAATGATTTAGAAATGGAGTTTGTCCAAGGATGGTAGCTGTATGT	1162
QY	1079	ATTTAACTCGGGCAAGAAATTTGTGAAAGAGTTTGTGTAGAAACCTGATGTTCTGAACCAA	1138
Db	1163	TGTTCACTACCGAAGAAAGAAATGAAAGAAATCGTGAGAAATCTATTTTGTGGAAGCAA	1222
QY	1139	ATGGGCAAAATGAGGGAACACATACCAACCACTGAAGAGCATGATCCAGTTTGTAAATCATTC	1198
Db	1223	ATGCTTCAATGAAGGATATATGCGGCCAGTCTCTGAGTATCTTAGCAATGCAATGAGTAC	1282
QY	1199	TGGGGGTGCTAACTCTGCTTACACAACTTGTTATCTTTGGCATGAGTGAGTANATATTTCAAA	1258
Db	1283	CAGCACATTTATCTTGCTTAGCATCACTATTTTGGCATCAAGTCTCTCAACAGCA	1342

Qy	1259	AGAGTCTGTGCGAATGGGCTGTCTCTGCACTCTCTCTTTTATAGATACATCAGGTATCTTGG	1311
Db	1343	AGATTTTGG---AATGGTTGGCCAAAGAACCCCTAAAAAATTCCTTGGAGCTAAATGTGACGTTATG	1399
Qy	1319	TCGAGCGCTAAATGACTCATGACCCACACAGGCCGAGCAGAAAGAAAGAAAAACATAGTTTCATC	1378
Db	1400	CCGAGTCATAGATGACATAGCCACCTATGAGGTTGAGAGGGTAGAGGTGAGATTTGCCAC	1459
Qy	1379	GAGCCTTGGAAAGTTATATGAAGGAATATATATGTCAATGAGGAGTATGCCCAACCTTGAT	1438
Db	1460	TGGAATTTGAATGTTTACATGAGAGATTATGCTGTATCCACAGAAAGGCCATGGAAAAATT	1519
Qy	1439	TTACAAGGAAGTAGAGATGTGTGGAAAGATATAAACCAGAGAGTACTCTACACTAAAAA	1498
Db	1520	CCAAGAAATGGCTGAGACAGCATGGAAGGATGTAAATGAAGGAATCCTTCGACCAACTCC	1579
Qy	1499	CATTCCAAAGCCGCTTATTGATGGCTGTGATCTATTGTTGCGAGTTCTTGGAACTTCAATA	1558
Db	1580	CGTCTCTACAGAGATTCTCTACTCGCAATCTCAATCTGCTCGCAATATCGATGTTACTTA	1639
Qy	1559	TGCAGGAA---AGGATAAATTCACACGTATGGGAGACGAATACAAACATCTCTATAAGTC	1615
Db	1640	TAAAGCAATCAAGATGGATACACTCATCCGAAAAAAGTACTAAAAACCTCATATTATTGC	1699
Qy	1616	TCCTACTCGTTATTCCTATGAGTATATGA	1643
Db	1700	GTTGTTGGTGGACTCTATTGAAATTTAA	1727
RESULT 9			
US-09-887-586A-31			
; Sequence 31, Application US/09887586A			
; Patent No. 6495354			
; GENERAL INFORMATION:			
; APPLICANT: Chappell, Joseph			
; APPLICANT: No. 64953541, Joseph P.			
; APPLICANT: Starks, Courtney M.			
; APPLICANT: Manna, Kathleen R.			
; TITLE OF INVENTION: SYNTHASE			
; FILE REFERENCE: 07678-025001			
; CURRENT APPLICATION NUMBER: US/09/887,586A			
; CURRENT FILING DATE: 2001-06-22			
; PRIOR APPLICATION NUMBER: 09/398,395			
; PRIOR FILING DATE: 1999-09-17			
; PRIOR APPLICATION NUMBER: 60/130,628			
; PRIOR FILING DATE: 1999-04-22			
; PRIOR APPLICATION NUMBER: 60/150,262			
; PRIOR FILING DATE: 1999-08-23			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 31			
; LENGTH: 1944			
; TYPE: DNA			
; ORGANISM: Solanum tuberosum			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (57)...(1724)			
; OTHER INFORMATION: vetispiradiene synthase			
US-09-887-586A-31			
Query Match 20.0%; Score 330.4; DB 4; Length 1944;			
Best Local Similarity 52.0%; Pred. No. 1.5e-87;			
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;			
Qy	14	AGAGAAAAACCTATTTCGGGCCCATTCGCCAACTTCTCCACAGATTTCGGGAGATCAGTT	73
Db	92	AGAGGAGGAGATTGTTTCGGCCCCCATAGCTGACTCTCTCCAAAGTCTTTGGGGGTATCGTTT	151
Qy	74	TCCTCATCTATCAAAACCAAGTAGACGACGGGTGGGAACAGATAGTGAATGATTTAAAAAA	133
Db	152	TCATTCAATTCCTCCGCAATCAGATTGCTGGAANAATATGCTCAAGAGATCGAAACTTT	211
Qy	134	AGAAGTCGCGCACTACTAAAAAGAACCTTTGGATATTTCTTATGAAACATGCCCAATTTGTT	193

Query Match	20.0%;	Score 330.4;	DB 4;	Length 1944;
Best Local Similarity	52.0%;	Prod. No. 1.5e-87;		
Matches 857;	Conservative	0;	Mismatches 761;	Indels 30;
				Gaps 4;

Qy	14	AGAGAAAAACCTATTGGCCCATTCCTCCAACTTTCTCCAGCATTTGGGAGATCAGTT	73
Db	92	AGAGGAGGAGATTGTTGGCCCATAGCTACCTCTCTCCAAGCTTTGGGGTGATCGTTT	151
Qy	74	TCTCATCTCAAAAGCAAGTAGAGCAAGGGTGGACAGATAGTGAATGATTTAAAAA	133
Db	152	TCATTCATTCTCCCTCGCAATCAGACTGCTGAAAATATGCTCAAGAGATCGAACTTT	211
Qy	134	AGAAGTGGCGCACTACTAAAGAAAGCTTTGGATATTCCTATGAACATGCCAAATTTGTT	193

Db 212 GAAGGAACAATCAAGAAATATATATCTGCACTCTTCGGAAGAACATTTGGCTGAGAAATT 271
QY 194 GAAGCTGATGTAGTAAATTCACGCGCTTGGAAATACCGTATCACATTTTGAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGTCAGCGCTTGGCAATTTGCTTATCATTTTGAACCAAAATAGA 331
QY 254 TCATGCAATTCGAATGATTTTATGAACATATGGT-----GATACTGGAA 298
Db 332 TGAATATGTTGGATCAATTTTCAAGCAGATCCCTAACTTTGAGGCTCACGAGTACAATGA 391
QY 299 TGGTGACCGCTCTCTCTATGGTTCCGCTTTATGCGAAAGCAAGGATATATGTTATGATG 358
Db 392 TTTACAACTTTATCCGTTCAATTTGACATTTGAGACACATGGTTACAAATATCTCCCC 451
QY 359 TGAATGTTTCAATPAACATATAAGACAAAATGAGCGCTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAACCTTTTATAGATTCACAGATGCAAAAGGCAAAATTTAAAGAACTCTCTTTGTAACGA 511
QY 419 TGTGTAAGGTTTCTTTGAGTTGACGAGCAACTTCTATGAGGTAAGTCTGGGAGATAT 478
Db 512 CATCAAGGCTCTTTTGAACCTTATACGAAGCCTCGCATGTAAGGACTCATGGAGAAGATAT 571
QY 479 ATTAGAAGATGCTCTCTGTTTACACGATCTCGTCTTAGCATATGACAAAAGATGCTTT 538
Db 572 TTTGAGAGGCACTTGTCTTCTACTGCTCACTTGAATCT-----GCAGCTCC 622
QY 539 TTTCTACAAACCCCGCTCTTTTTCACGAAATPACAAACGGGCACTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCTTGAGCAATCTCTCCATAA 682
QY 599 AAGGTTGCCAAGATAGAGCGCGCGAGTACATCTCTTCTATCAACAAAGATTTCTCA 658
Db 683 GAGCATTTCCAAGAGTTGAGACACGCTACTTCTATCTATCTACGAAGAGGAGAACAGAA 742
QY 659 TAACAGACTTTTACTTAACTTCTAGTTAGAGTTCAATTTGCTTCACTGATTCACAA 718
Db 743 GAATGATGTTGCTTCAATTTCCAAACCTGGACTTCACTTCACTTCACTGATTTGCAAA 802
QY 719 GGAAGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCATATCAAGAAAGCAGCAC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGGTGGAAAGATTTGGAATTTGTGACACACATCTC 862
QY 779 TGTGTTAAGATAGATATGTTGAATGCTACTTTTGGGAGCTAGGTTTCAGGCTATGAGCC 838
Db 863 ATATGCTAGGGATAGACAGTGGAAATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
QY 839 ACAGTATTCGGGCTAGAGTTTCTTTCACAAAGCTGTTGCTTTAATACTCTTATAGA 898
Db 923 TCAATACCTCAGGCTCGTGTGATGCTTGAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TCACACTTATGATGCGTATGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTTGA 958
Db 983 TGACACATTCGATGCTTATGGCATTTGCAAGAACTTGAGATCTACACCGATGCCATACA 1042
QY 959 AAGGTGGTCAATATCATGCTTAGACACATCTCAGAAATACATGAACCGGATATACAAAT 1018
Db 1043 GAGGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATATACATGAATATCAGTATCAAA 1102
QY 1019 ATTATGATATACATACACAGAAATGGAAGATTTCTTGAAGAGGAGGGAAGACAGATCT 1078
Db 1103 ACTTTAGATCTCAATGATATGAAATGGAAGTTGTCCAGAGTGTCCAGAGTGGTAGATCTGAT 1162
QY 1079 ATTAACTCGGCAAGAAATTTGTAAGAGTTTGTGAAGAACTGATGTTGTAAGCAAA 1138
Db 1163 TGTCTACTACCGGAAAGAAAGAAATGAAGAAATCGTGAGAAATCTATTTTGTGAAGCAAA 1222
QY 1139 ATGGGCAATGAGGGACATACACCACTGAGAGCATGATCCAGTTGTAATCATATC 1198
Db 1223 ATGGTTTATTGAAGGATATATGCCCGCAGCTCTGAGTATCTTAGCAATGATATGATCTAC 1282
QY 1199 TGGCGGTGCTAACCTCTTACAACTCTGTTATCTTGGCATGAGTATATTTTCAAAA 1258

Db 1283 CAGCACTTATTACTTGTCTTACGACTACATCTTATTTGGGCATGAAGTCTGCTAACAGCA 1342
QY 1259 AGAGTCTGCAATGGGCTGTCTCTGCACTCTCTCTTTTAGATACTCAGGTATCTTGG 1318
Db 1343 AGATTTTG--AATGTTGGCCCAAGAACCTTAAATTTCTTGGGCTAAJGTGACGTTATG 1399
QY 1319 TCGAGCCCTAAATGATCTCATGACCCCAAGCCCGAGGAGCAAGAAACATATGTTTCATC 1378
Db 1400 CCGAGTCTATAGATGATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
QY 1379 GAGCCTTCAAAAGTTATATGAAGGAATATAATCTCAATGAGGAGTATGCCCAACCTTGTAT 1438
Db 1460 TGAATTTGAATTTTACATGAGAGATTTGTTGATATCCACAGAAAGGCCATGGAATAAT 1519
QY 1439 TTACAAGGAAGTAGAAGATGTTTGGAAAGATATAAACCAGAGATGATCTTCAAACTAAAAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGATGTAATGAAGAAATCTCTTCGACCACTCC 1579
QY 1499 CATTCGAAGGCGTTTATGATGGCTGTGATCTATTTTGGCCAGTTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTTCTCACTCGCATTTCTCAATCTTCTGTCATTTATCGATTTACTTA 1639
QY 1559 TGCAGGAA--AGGATAACTTTCACACGATGATGGAGACGAATACAAACATCTCATAAGTC 1615
Db 1640 TAAGCAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1699
QY 1616 TCTACTCGTTTATCCTATGATGATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 10

US-09-895-752-31
; Sequence 31, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895, 752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-895-752-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Pred. No. 1.5e-87;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;
QY 14 AGAAGAAAAACCTATTCGCCCATTTGCCAATTTCTTCTCCAGCATTTGGGAGATCAGTT 73
Db 92 AGAGGAGAGATTTGTCGCCCATAGCTGACTTCTCTCCAGTCTTTGGGAGATCGTTT 151
QY 74 TCTCATCTATAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATTTTAAAAA 133

Db 152 TCATTCTCCCTCGCAATCAGATTGCTGGAATAATGCTCAAGAGATCGAACTTT 211
QY 134 AGAGTCGGGCAACTACTAAAGAAGCTTTGGATATTCCTAAGAAACATGCGCAATTTGTT 193
Db 212 GAAGGAACAAATCAAGAATATATATCTGCAATCTTCTCGAAGAAATATGGCTGAGAAAT 271
QY 194 GAAGCTGATTGATGAAATTCAGGCTTTGGAATACCGTATCACTTTGAAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGAGGCGCTTGGCAATGCTTATCATTTGAAACAAATAGA 331
QY 254 TCATGCAATTGCAATGATTTATTAAGAAATATGCTT-----GATAACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTACAAAGCAGATCCTAACTTTGAGGCTCAGGAGTACAAATGA 391
QY 299 TGGTGACCGCTCTTCCTTATGGTTCGCTCTTATGCGAAAGCAGAGATATATGTTACATG 358
Db 392 TTTACAAACTTTATCCGTTCAATTTGAGCTATTTGAGCAACATGTTTCAATATCTCCCC 451
QY 359 TGATGTTTTCATTAATATAAGACAAATAATGAGCGTTTCAAGCAATCGTTAGCTAAATGA 418
Db 452 AARACTTTTATTAGATTCGAAGTGCRAAGGCAATTTAAGAAATCTCTTTGTAAGA 511
QY 419 TGTGAAGTTTGTGATTTGATAGAGCAACTTCTATGAGGGTACCTGGGAGATTAT 478
Db 512 CATCAAGGCTCTTTGAACTTATACGAAGCTCGCATGTAAGGACTCATGAGAGAGATAT 571
QY 479 ATTAGAGATGCTCTTGGTTTTACAGATCTGCTTTAGCATATGACAAAGATGCTTT 538
Db 572 TTTGGAAGAGGCACTTGCTTTCTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
QY 539 TTCTFACAAACCCCGCTCTTTTACGGAATACAAAGCGCACTTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTGAGTCACTCTGAGTAGCAAGTAGACATGCCCCTTGAGCAATCTCTCCATAA 682
QY 599 AAGTTGCCAAGATAGAGGGGGGCGAGTACATCTCTTTCTATCAACAAAGATTCTCA 658
Db 683 GAGCATTTCCAGAGTGTAGACACGCTACTTCTCTATCTACGAGAGGAGGACAGAA 742
QY 659 TAACAAGACTTTACTTAACTGCTTAAGTTAGATTCATTAATTTGCTTCACTAGTCAATTCACAA 718
Db 743 GAATGATGTGTTGTTCAATTTGCAAACTGGACTTCAACTTACTTCAGATGTTCCACAA 802
QY 719 GGAAGAGCTCAGGCATGTGTCAAATGCTGGAAGCTTTTCGATATCAAGAAGAACGCACC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTGGAAGATTGGAATTTGTGACAACTCTCC 862
QY 779 TTGTTTAAGAGATAGAAATTTGTTGAATGCTACTTTTGGGCACTAGGTTTCAAGCTATGAGCC 838
Db 863 ATATGCTAGGATAGACAGTGGATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
QY 839 ACAGTATCCCGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTAACTCTTTATAGA 898
Db 923 TCAATACTCTCAGGCTCGTGTGATGCTTGTCTAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATCGGTATGCTTATGAGAACTTAAGATCTTTACTGAAGCTTTGA 958
Db 983 TGACACTTATGCTTATGCTATGCTATGCTTCAAGAACTTGAATCTACACCGATGCCATACA 1042
QY 959 AAGTGTGCTAATTATACATGCTTAGACACACTTCCAGATATACATGAACCGATATACAAAT 1018
Db 1043 GAGTGGATATTAGCCAAATTTGATCGCTCCCTGATTTACATGAATACTAGTACAAAGC 1102
QY 1019 ATTCAATGATATACACAGAAATGGAAGAAATTTCTCAAGAGGAGGAAAGCAAGATCT 1078
Db 1103 ACTTTTAGATCTCTACAAATGATTAATGAATGGAGTTGTCCAAGGATGGTAGATCTGATGT 1162
QY 1079 ATTTAACTGGGCAAGAAATTTGGAAGAGTTTGTGGAAGCTTGTGGAAGCTTGTGAGCAAA 1138
Db 1163 TGTTCATACCGCAAGAAAGAAATGGAAGAAATCTGTGAGAACTATTTTGTGGAAGCAAA 1222
QY 1139 ATGGGCAATGAGGAGACACATACCAACCACTGGAAGAGATGATCCAGTTGTAATCATTAAC 1198

Db 1223 ATGTTTCATTGAAGGATATATGCGGCAGTCTCTGAGTATCTTAGCAATGCAATTAGTAC 1282
QY 1199 TGGCGGTGCTAACTGCTTACAACTTGTATCTTGGCATGAGTATATATTCACAAA 1258
Db 1283 CAGCACTTATTTACTTGGTCTAGCACTACATCTTATTTGGGCATGAGTCTGCTTAAACGCA 1342
QY 1259 AGAGTCTGTGCAATGGGCTGCTCTGCACTCTCTCTTTTATAGATCTCAGGTATACTTGG 1318
Db 1343 AGATTTTG---AATGTTGGCCAAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TGACGCCCTAAATGATCTCATGCCCCAACAGGCCGAGCAAGAAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACTATGAGTTGAGAGGGGTAGAGGTGAGATTGCCAC 1459
QY 1379 GAGCCTTGAAGTTATATGAGGAATATATGATCAATGAGGAGTATGCCAAAACCTTGAT 1438
Db 1460 TGGAAATGGAATGTTATACATGAGAGATTATGTTGATCCACAGAAAAGGCCATGGAATAAT 1519
QY 1439 TTAACAGGAAGTAGAAGATGTGTGGAAGATATAAACCAGAGAGTACCTCACAACTAAAAA 1498
Db 1520 CCAGAAATGGCTGAGACAGCATGGAAGGATGTAAATGAAGGAATCCTTCGACCACTCC 1579
QY 1499 CATTCGAAGCCGTTATTTGATGCTGTGATCTATTTTGGCAGTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTCTCACTCGCATTTCTCAATCTTGTCTGCGATTATCGATGTTACTTA 1639
QY 1559 TGCAGGAA---AGGATAACTTCACAGTATGGGAGAGCAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCAAAATCAAGATGGATACACTCATCCGAAAAAGTACTTAAACCTCATATTATTGC 1699
QY 1616 TCTACTCGTTTATCTCTATGATATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 11
US-09-903-012B-31
; Sequence 31, Application US/09903012B
; Patent No. 6589656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65896561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 50/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-903-012B-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Pred. No. 1.5e-87;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;
QY 14 AGAAGAAAAACCTATTTCGCCCACTTCCCAAGCAATTTCCCAAGCAATTTGGGAGATCAGTT 73
Db 92 AGAGGAGGAGATGTTTCGCCCACTAGTCTCTCCAAAGTCTTTGGGAGTATCGTTT 151

QY 74 TCTCATCTATCAAAAGCAAGTAGCAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTCATCTCTCCCTCGACAAATCAGATGCTGGAAATATATGCTCAAGAGATCGAAATCTTT 211
QY 134 AGAAGTCGGCGCACTACTAAAAAGAGCTTTGGATATTCCTATGAAACATGCCAATTTGTT 193
Db 212 GAAGGAACAATCAAGAAATATATATTCATCTCTCGAAGAACATTTGCTGAGAAAT 271
QY 194 GAAGCTGATGATGAATTAACAGCCCTTGGAAATACCGTATCACTTTGAACGGGAGATGGA 253
Db 272 GGATCTGATAGACATTTGTTGAGCGCTTGGCAATGCTTATCATTTTGAAGAAACAAATAGA 331
QY 254 TCATGCAATTCGAATGATTTATGAAACATATGTT-----GATTAACGGAA 298
Db 332 TGATATGTTGGATCAATTTTAAAGCAGATCTTAACCTTTGAGGCTCACGAGTACAAATGA 391
QY 299 TGGTGAACCGCTCTTCCTATGTTTCCGTCTTATGCGAAAGCAAGGATATATGTTACATG 358
Db 392 TTTTCAAACTTTTATCCGTTCAATTTTCGACTATTGAGACAAATGTTTACAATATCTCCCC 451
QY 359 TGATGTTTCAATACTATTAAGACAAATAATGAGCGTTCAAGCAATCGTTAGCTTAATGA 418
Db 452 AAAACCTTTTATTAGATTCGAAGATCCAAAGAGCAAAATTTAAAGAAATCTCTTTGTAACGA 511
QY 419 TGTTGAAGGTTTGTGAGTGTGACGAAGCAACTTCTATGAGGTTACCTGGGGAGATTAT 478
Db 512 CATCAAGGTCCTTTGAACTTATACGAAGCTCGATGTAAAGACTCATGGAGAAATAT 571
QY 479 ATTAGAAGTGTCTTTGGTTTACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTT 538
Db 572 TTTGGAAGAGGCACTTGTCTTCTACTGCTCATCTTGAATCT-----CGAGCTCC 622
QY 539 TTTTACAAACCCGCTCTTTTACCGAAATPACAAACGGGCACTAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCTTGAGCAATCTCTCCATAA 682
QY 599 AAGGTTGCCAAGATAGAGCGCGCGAGTACATCTCTTCTATCAACAAAGATCTTCA 658
Db 683 GAGCATTCGAAGTTGAGACAGCTACTCTATCTATCTAGAGAGAGAGAAAGAA 742
QY 659 TAACAAGACTTTTACTTAAATCTGCTAAGTTAGAGTTCAATTTGCTTCACTTCACTGACAA 718
Db 743 GAATGATGTTGCTTCAATTTGCAAACTGGACTTCACTTACTTCACTTCACTTCACTTCA 802
QY 719 GGAAGAGCTCAGCATGCTGCAAAATGTTGGAAAGCTTTGATATCAAGAAAGCGAAC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTTGGAAGATTTGGAATTTGTGCAACACTTCC 862
QY 779 TTGTTTAAAGATAGAAATTTGAATGCTACTTTTGGGAGTACTAGTTTCAGGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGTGTATGCTGAACC 922
QY 839 ACAGTATTCGGGGCTAGAGTTTCTTCAAAAGCTTTGCTTATTAATCTCTTATAGA 898
Db 923 TCAATACTCTCAGGCTCGTGTATGCTTGTCTAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATGCTATGCTTATGATGAGAACTTAAGATCTTTTACTGAAAGCTGTTGA 958
Db 983 TGACACATTCGATGCTTATGCTATTCGAAGAACTTTGATCTACACCGATGCCATACA 1042
QY 959 AAGGTTGCTCAATTTACATGCTTACACACACTTCCAGAAATACATGAACCGATATACAAAT 1018
Db 1043 GAGTGGGATATAGCCAAATTTGATCGGCTCTGATTTATCATGAATCATGTTTACAAAGC 1102
QY 1019 ATTCATGATACATACAGAAATGGAAGAAATTTCTTGAAGAGGAGGAAAGACAGATCT 1078
Db 1103 ACTTTTAGATCTCAATGATTTATGAATGAGTGTGTCACAGATGTTAGATCTGATGT 1162
QY 1079 ATTTAACTCGGCAAGAAATTTGTGAAAGTGTGTAAGAACTGATGTTTGAAGCAAA 1138
Db 1163 TGTTCACTACGCGAAAGAAAGATGAAGAAATCGTGAGAAATCTATTTTGTGGAAGCAAA 1222

QY 1139 ATGGGCAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATCATTTAC 1198
Db 1223 ATGTTTCATTTGAAGATATATGCCCCAGTCTCTGAGTATCTTAGCAATGCAATTTAGCTAC 1282
QY 1199 TGGCGGTCTAAACCTGCTTACAAACAACTTGTATCTTGGCATGAGTATATTTCAAA 1258
Db 1283 CAGCACTTATTTACTTGTCTTACGACTACATCTTATTTGGCATGAAAGTCTGCTAAACAAGA 1342
QY 1259 AGATCTGTGGAATGGGCTGTCTGACCTCTCTTTTAGATATCTCAGGTATATCTGG 1318
Db 1343 AGATTTTG---AATGGTTGGCCAAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TCGAGCGCTTAATGATCTCATGACCAAGCGCAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGTACATAGCCACTATGAGGTTGAGAGGGTAGAGTCAAGTCTCCAC 1459
QY 1379 GAGCTTTGAAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCAACCTTGAT 1438
Db 1460 TGAATTTGAATGTTTACATGAGAGATTATGGTGTATCCACAGAAAGGCCATCGAAAAAT 1519
QY 1439 TTACAAGGAAGTAGAAGATGTTGGAAAGATATAAAACGAGAGTACCTCACAACATAAAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATTAATGAAGGAATCCCTTCGACCAACTCC 1579
QY 1499 CATTCGAAGCGCTTATTTGATGGCTGTATGATCTTATTTGCCCAGTTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATCTCTACTCGCATCTCAATCTTCTGCTCCATTTATCGATTTACTTA 1639
QY 1559 TGCAGGAA---AGGATAACTTTACACGATGATGGAGAGCAATAACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCACAATCAAGATGATACATCTCATCGGAAAAAGTACTTAAACCTCATATTTATTCG 1699
QY 1616 TCTACTCGTTTATCTCTATGATATGA 1643
Db 1700 GTTGTGCTGACTCTATTGAAATTTAA 1727

RESULT 12
US-09-900-797-31
; Sequence 31, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6645762L, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-900-797-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Fred. No. 1.5e-87;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;
QY 14 AGAAGAAAAACCTATTGCCCCCAATTCGCAACTTTCTCTCAAGCAATTTGGGAGATCAGTT 73

Db 92 AGAGGAGGATTTGTCGCCCCATAGCTGACTTCTCTCAAGTCTTTGGGGTGATCGTTT 151
QY 74 TCTCATCTATCAAAAGCAAGTAGACCAAGGGGTGGAAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCAATTCATCTCCCTCGACAATCAATTCGTCGAAAATATGCTCAAGAGATCGAAACTTT 211
QY 134 AGAAGTGGGCAACTACTATAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTT 193
Db 212 GAAGGAACAATCAAGAAATTAATTAATCTGCAATCTCTCGAAGCAATTTGGCTTGAGAAAT 271
QY 194 GAAGCTGATGATGAATTAACAGCTTTGGAAATACCGTATCACTTTTGAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGTTGAGCGCTTGGCATTCCTTATCATTTTGAAGAAACAATAGA 331
QY 254 TCATGCAATGCAATGATTAATTAAGAAATATAGGT-----GATAACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTCAAGAGCAGATCCTAACTTTGAGGCTCAGAGTACAATGA 391
QY 299 TGGTGACCGCTCTTCTTATGTTGCTTCCGCTTATGCGAAGCAAGGATATTTATGTTACATG 358
Db 392 TTTACAAACTTTATCCGTTCAATTTTGGATTTTGAACAACATGTTTACAATATCTCCCC 451
QY 359 TGATGTTTTCAATTAACATATAAGACAAATAATGGAGCGTTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAACTTTTATTAGATTCGAAGATGCAAGAGGCAAAATTTAAGAAATCTCTTTGTAACA 511
QY 419 TGTGAAGTTTGGCTTGTGATTTGAGAGCAACCTCTATGAGGGTACCTGGGAGATTAAT 478
Db 512 CATCAAGGGTCTTTTGAATTTATCAAGAGCCTCGCATGTAAGGACTCATGGAAGAGATAT 571
QY 479 ATTAGAAGATCTCTTGTGTTTTACAGCATCTGCTTTAGCATTAATGACAAAGAGTCTTT 538
Db 572 TTTGGAAGAGCACTTGTCTTCTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
QY 539 TTCTACAAACCCCGCTCTTTTTACGAAATACAGGGCAGCTTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAGTCACTCTGAGTAAAGCAAGTACATGCGCTTGAAGCAATCTCTCCATTA 682
QY 599 AAGTTGCCAAGAAATAGAGGGGGGCGAGTCAATCTCTTCTATCAACAAAGATCTCA 658
Db 683 GAGCAATTCGAAGAGTTGAGACAGCTACTTCTATCTCTATCTACGAAGAGGAGGAACAGAA 742
QY 659 TAACAAGACTTTACTTAACCTTCTAGTGTAGAGTTCAATTTGCTTCAGTCAATTCACAA 718
Db 743 GAATGATGTTGCTTCAATTTGCAAACTGGACTTCACTTACCTTACAGATGTTGCACAA 802
QY 719 GGAAGAGCTCAGCCATGTCGCAATGTTGGAAGCTTTTCGATATCAAGAAAGCGCACC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTTGGAAGATTTGGATTTTGACACACTTCC 862
QY 779 TTGTTTAAAGATAGAAATGTTGAATGCTACTTTTGGGAGCTAGGTTTCAAGGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGGTGTATGCTGAACC 922
QY 839 ACAGTATTCGGGCTAGAGTTTCTTACAAAGCTGTTGCTGTTATTAACCTTTATAGA 898
Db 923 TCAATCTCTCAGGCTCGTGCATGCTTGTCTTAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATCGGTATGCTATTAAGAACTTAAAGATCTTTTACTGAAGCTGTTGA 958
Db 983 TGACACATTCAGUCTTATGCAATTTGCAAGAACTTGAGATCTACCGATGCCATACA 1042
QY 959 AAGTGTGCAATTACATGCTTTAGACACACTTCCAGAAATACATGAACCCGATATACAAAT 1018
Db 1043 GAGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATTACATGAATACTAGTTACAAAGC 1102
QY 1019 ATTCAATGATACATACAGAAATGGAAGAAATTTCTTCAAGAGGAGGAGNAGACAGATCT 1078
Db 1103 ACTTTTATGCTCTACAAATGATTAAGAAATGGAGTTGTCCAAAGGATGGTAGATCTGATGT 1162
QY 1079 ATTTAACTGGCGCAAGAAATTTGTGAAGAGTTTGTGTAAGAACTGATGTTTGAAGCAAA 1138

Db 1163 TGTTCACTAGCGAAAGAAAGAAATGAAAGAAATCGTGCAGAAACTATTTTGTGGAAGCAAA 1222
QY 1139 ATGGGCAATAGAGGACACATATCAACCACTGAAGACATGATCCAGTTTGTAAATCATTTAC 1198
Db 1223 ATGGTTTCAATTTGAAGGATATATGCGCCAGTCTCTGAGTATCTTAGCAATGCAATTAGCTAC 1282
QY 1199 TGCGCGTGTCTAACTGCTTACAACTTGTATCTTTGGCATGAGTGATATATTTACAAA 1258
Db 1283 CAGCACTTATTAATCTGCTTACGACTACATCTTATTTGGGCATGAAGTCTGCTAAACAGCA 1342
QY 1259 AGAGTCTGTGAATGGGCTGCTCTGTCACCTCTCTTTTATAGATACTCAGGTATACTTGG 1318
Db 1343 AGATTTTG--AATGGTTGGCCAAAGAACCTTAAAAATCTTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TGGACCCCTAAATGATCTCATGACCCCAAGGCCGACGAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACTATGAGTTTGAAGGGTAGAGGTCAAGTTGCCAC 1459
QY 1379 GAGCCTTTGAAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCCAACCTTGAT 1438
Db 1460 TGGAAATGAAATGTTATCATGAGAGATTTATGTTATCCACAGAAAGGCCATGGAAAAAT 1519
QY 1439 TTACAAGGAATGAAGATGTTGGAAAGATATAAACCAGAGAGTACCTCAACAATAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATGTAATGAAGGAATCTTTCGACCAACTCC 1579
QY 1499 CATTCGAAGCCGTTATGATGCTGTGATCTATTTGTGCCAGTTTCTTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTTCTCACTCGCATTTCTCAATCTTTGCTGCGATTATCGATGTTACTTA 1639
QY 1559 TGCAGGAA--AGGATAACTTCAACGTATGGGAGACGATACAAACATCTCATAAAAGTC 1615
Db 1640 TAAGCAATCAAGATGGATACACTCATCCGAAAGAAAGTACTAAACCTCATATTTATTGTC 1699
QY 1616 TCTACTCGTTTATCTCTATGAGTATAGA 1643
Db 1700 GTTGTGGTGGACTCTTATTGAATTTAA 1727

RESULT 13
US-09-398-395A-1
; Sequence 1, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)....(1668)
US-09-398-395A-1

Query Match 19.4%; Score 320.6; DB 4; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.1e-84;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAACCTATTTCGCCCACTTCCCACTTTCTCCCAAGCATTTGGGAGATCAGTTT 74
Db |||||
QY 52 GAAGAAAGATTTGTTGGCCCGTCCGAGCTTCTCCCTAGTTTGGGGGTGATCAGTTT 111
Db |||||
QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAAA 134
Db |||||
QY 112 CTTTCATTTCCATTAA ---AAATCAGGTTGCAGAAAAGTAGTCTCAAGAGATTGAAGCA 168
Db |||||
QY 135 GAAGTGGCGCAACTACTAAAGAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTG 194
Db |||||
QY 169 TTGAAGGAACAAACGAGGAATATGCTGTTAGCAACTGGAATGAAATTTGGCTGATACACTG 228
Db |||||
QY 195 AAGCTGATTGATGAATTAACAGCTTGAATACCGTATCATCTTCAACGGGAGATTGAT 254
Db |||||
QY 229 AATTGATGACACTATTGAACGCTTTGGCATATCTCCACCTTTGAGAAAGAAATGAT 288
Db |||||
QY 255 CATGCATTGCAATGATTTATGAACATATGTTGATPACTGGAATGGTGACCGCTCTTC 314
Db |||||
QY 289 GATATTTTGGATCAGATTTACACCAAACTCAAACTGCAACGATTTGTGCACCTCTGCA 348
Db |||||
QY 315 TTATGGTCCGCTTATGCGAAAGCAAGATATATGTTACATGATGTTTCAATTAAC 374
Db |||||
QY 349 CTTCAATTTGATGCTGCTCAGGCAACATGTTTCAACATCTCTCCTGAAATTTTCAGCAA 408
Db |||||
QY 375 TATTAAGACAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCCT 434
Db |||||
QY 409 TTCCAAGAGAAAATGGCAAAATCAAGGAATCTCTTGTAGTATGTTCTTAGAATTTG 468
Db |||||
QY 435 GAGTTGTACGAAGCACTTCTATGAGGTACCTGGGGAGATTAATTAGAAGATGCTCTT 494
Db |||||
QY 469 AACTTGTATGAAGCTTCAATGTPAAGSACTCATGCTGACGATATCTTAGAAGCGCACTT 528
Db |||||
QY 495 GGTTTTACAGATCTCGCTTAGCATATGACAAAAGATGCTTTTCTCAAAACCCGCT 554
Db |||||
QY 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCACTCCACATTTGAAATCTCCA 579
Db |||||
QY 555 CTTTTTACGGAATAACACGGGCATTAAGCAACCCCTTTGGAAGAGTTGCCAAGATA 614
Db |||||
QY 580 CTTAGGAGCAAGTGACATGCCCCCTTGAGCAATGTTTGACAGGGTGTCTTAGAGTC 639
Db |||||
QY 615 GAGGCGCGCAGTACATTCCT---TCTATCAACAACAGATTTCTCATACAAGACTTTA 671
Db |||||
QY 640 GAGACCGGATCTTCTCATCTCATCAATCTATGACAGGAACAATCGAAGATAATGTGTA 699
Db |||||
QY 672 CTTAAACTGTGTAAGTAGATTTCAATTTGCTTCAGTCAATGACAAAGGAGAGCTCAGC 731
Db |||||
QY 700 CTTGATTTGCCAAATGGAATTTCAACTTCTCCAGATGTTGACAAACAAAGAACTTGCT 759
Db |||||
QY 732 CATGCTGCAAAATGGTGGAAAGCTTTGATATCAAGAAAGACGACCTTGTTTAAGAGAT 791
Db |||||
QY 760 CAAGTATCAAGGTGGTGAAGATTTGGATTTTGTATCAACACTTCCATATGCTAGAGAT 819
Db |||||
QY 792 AGAATTTGTTGAATGCTACTTTTGGGACTAGTTTCAAGCTATGAGCCACAGTATTTCCCGG 851
Db |||||
QY 820 CGAGTAGTTGAATGCTACTTTTGGSCATTAGGAGTTTATTTGAGCCTCAATCTCTCA 879
Db |||||
QY 852 GCTAGAGTTTCTTCAAAAGCTGTTGCTGTTATAACTCTTATAGATGACATTTATGAT 911
Db |||||
QY 880 GCTCGGCTGATGCTCGTTAAGACCATATCAATGATTTTCGATTTGCGATGACACCTTTGAT 939
Db |||||
QY 912 GCGTATGCTACTTATGAGAGACTTAGATCTTTACTGAAGCTGTTGAAGGTGCTCAATT 971
Db |||||
QY 940 GCTTACGGTACATTTAAGAACTTGAGGCATACACAGATGCCATACAAAGATGGGATATC 999
Db |||||
QY 972 ACATGCTTAGACACACTTCCAGAAATACATGAACCCGATATACAAATATTATTCAGTATACA 1031
Db |||||
QY 1000 AACGAATTTGATCGGCTTCTCGATTACATGAAATCAGTTACAAAGCTATTTCTAGATCTC 1059
Db |||||
QY 1032 TACACGAATGAAGAAATTTCTTTGCAAGAGGGAAGAAAGACAGATCTATTAACTGGCGG 1091
Db |||||
QY 1060 TACAAGGATTTATGAAGGAATTTGCTAGTGGCGGAAGATCTCATATTTGCTGCGCATGCA 1119
Db |||||
QY 1092 AAAGAAATTTGTAAGAGATTTGTTAGAAACCTGATGTTGAGCAAAATGGGCAATGAG 1151
Db |||||

Db 1120 ATAGAAGAATGAAGAAGTAGTAAGAAATTTATATGTCAGTCAACATGTTTATTGAA 1179
|||
QY 1152 GGACACATACCAACCTCTGAAGAGATGATCCAGTTGTAATCATTTCTGGCGGTGCTAAC 1211
|||
Db 1180 GGATATACGCCACTGTTTCTGTAATCCTAAGCAATGACCTAGCAACTACCAATATTAC 1239
|||
QY 1212 CTGCTTACACAACTGTTTATCTTTGGCATGAGTGATATATTACAAAAAGAGTCTGTCGAA 1271
|||
Db 1240 TACCTCGGCAACATCGTATTTGGGCATGAAGTCTGC---ACGGAGCAAGATTTTGAG 1296
|||
QY 1272 TGGGCTGCTCTGCACCTCTCTTTTTTAGATACTCAGGTATCTTTGGTCGACGCCCTAAAT 1331
|||
Db 1297 TGGTTGTCAAGAATCCAAAAATTTCTTGAAGCTAGTGTAAATATATGTCGAGTTATCGAT 1356
|||
QY 1332 GATCTCATGACCCCAAGCGCGGAGCAAGAAAGAAACATAGTTTCTATCGAGCTTGAAGT 1391
|||
Db 1357 GACACACCCACGTACGAGGTTGAAAAAGCGAGGGGACAAATTTGCAACTGGAAATGAGTGC 1416
|||
QY 1392 TATATGAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAAGGAAGTA 1451
|||
Db 1417 TGCATGAGATTAATGTTATATCAACAAAAGAGGCAATGGCTAAATTTCAAAATATGGGT 1476
|||
QY 1452 GAAGATGTGTGGAAGATATAAACCGGAGAGTACCTCAACACTTAAAAAATTTCCAAAGCCG 1511
|||
Db 1477 GAGACACATGGAAGAATATTAATGAAGGACTTCTTAGGCCCACTCCCGTCTCTACAGAA 1536
|||
QY 1512 TTATGATGGCTGTGATCTATTGTGCCAGTTTCTTGAAGTTCAATATGACAGAAA--G 1568
|||
Db 1537 TTTTAACTCTCTATTTCTCAATCTTGTGCTGATTTGTGAGGTTACATATACAAATCTA 1596
|||
QY 1569 GATAACTTCAACGCTATGGGAGAGCATACAAAACATCTCATAAAGTCTCTACTCTGTTTAT 1628
|||
Db 1597 GATGGATACACTATCGGAGAAAGTCTTAAACCTCAATTTAACTTAACTTGTGGAC 1656
|||
QY 1629 CCTATGATATATGA 1643
|||
Db 1657 TCCATCAAAATTTGA 1671
|||

RESULT 14
US-09-887-586A-1
; Sequence 1, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887, 586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398, 395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130, 628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150, 262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)...(1668)
US-09-887-586A-1

Query Match 19.4%; Score 320.6; DB 4; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.1e-84;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAACCTATTGCGCCCAATTCGCCAATTCCTCCCAAGCATTTGGGGAGATCAGTTT 74
 Db 52 GAAGAAGAGATTGTTGGCCCGTCCGCCACTTCTCCCTAGTTTGGGGGTGATCAGTTT 111
 QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAABAA 134
 Db 112 CTTTCATTTCCATTAA---AAATCAGGTTCAGAAAGATATGCTCAAGAGATTGAAGA 168
 QY 135 GAAGTGGCGCACTACTAAAGAAAGCTTTGGATATTCCTATGAACATGCCAATTTGTTG 194
 Db 169 TTGAAGGAACAAACGAGGAATATGCTTTAGCAACTGGAATGAAATTTGCTGATACAC 228
 QY 195 AAGCTGATGATGAATCAAGCCCTTGAATACCGTATCACTTTGAACGGAGATTGAT 254
 Db 229 AATTGTATAGACACTAATTGAACCGCTTGGCATATCTTACCACTTTGAGAAAGAAATTGAT 288
 QY 255 CATGCATTTGCAATGATTTATGAACATATGTTGATTAACCTGGAATGGTGACCGCTCTCC 314
 Db 289 GATATTTTGGATCAGATTTAACCAAACTCAAACTGCAACGATTTGTGCACTTCTGCA 348
 QY 315 TTATGGTTCCGCTCTTATCGAAAGCAAGGATATTAATGTTACATGATGATGTTTCAATAAC 374
 Db 349 CTTCAAATTTGCTGCTCAGGCAACATGTTTCAACATCTCTCCTGAAATTTTCAGCAA 408
 QY 375 TATAAGACAAAATAGAGCGCTTCAAGCAATGTTAGCTATGATGTTGAAGTTTGCCT 434
 Db 409 TTCCAAGACGAAATGGCAATTTCAAGAAATCTCTTGTAGTGTCTTAGGATTTATG 468
 QY 435 GAGTTGTACGAAGCAACTTCTATAGGCTACCTGGGAGATATATAGAGATGCTCT 494
 Db 469 AACTTGTATGAGCTTCACTGTAAGGACTCATGCTGAGATATCTTAGAAGCGCACT 528
 QY 495 GGTTTTACCGATCTGCTTATGCAATATGCAAAAGATGCTTTTCTCAAAACCCCGCT 554
 Db 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCACTCCACATTTGAAATCTCCA 579
 QY 555 CTTTTCACGAAATACACGGGCACTAAGCAACCCCTTTGGAAGAGTTGCAAGATA 614
 Db 580 CTTAGGAGCAAGTAGACATGCTCCCTTGGCAATGTTTGCAACAGGGTGTCTTAGATC 639
 QY 615 GAGCGGGCAGTACATTCCTT---TCTATCAACCAAGATTCCTCAACCAAGCTTTA 671
 Db 640 GAGACCGGATTTCTCATCTCATCTATGCAAGGAAACATCGAAGATATGTTGTTA 699
 QY 672 CTTAAACTTGTAAAGTTAGATTCATTTGCTTCACTATTCGCAAGGAAGAGCTCAGC 731
 Db 700 CTTGATTTGCCAATTTGAATTTCACTTGTCTCAGATGTTGCAACCAAGCACTTGT 759
 QY 732 CATGTGCAAAATGTTGAAGCTTTGATATCAAGAAAGCGACCTTTGTTAAGAT 791
 Db 760 CAAGTATCAAGGTGGTGAAGAAATTTGATTTTGTAAACCACTTCCATATGCTAGAT 819
 QY 792 AGAATTTGTAATGCTACTTTTGGGAGCTAGGTTGAGGCTATGAGCCAGTATTCGCG 851
 Db 820 CGAGTAGTTGAATGCTACTTTTGGGCATTAGAGTTTATTTGAGCTCAATCTCTCA 879
 QY 852 GCTAGAGTTTCTTCAAAAAGCTGTGTGTTTATAACTCTTATAGATGACACTTATGAT 911
 Db 880 GCTCGCTCATGCTGTTAAGACCATATCAATGATTTGATTTGATGATGACACTTGT 939
 QY 912 CGGTATGTTACTTATGAAGACTTAAGATCTTTA CTGAAGCTGTTGAAGGGTGTCAAT 971
 Db 940 GCTTACGTTACAGTTTAAAGAACTTGAGGCATACAGATGCCATACAAAGATGGGATC 999
 QY 972 ACATGCTTAGACACACTCCAGATACATGAACCGATATACAAATTTATCTAGATACA 1031
 Db 1000 AACGAAATGATCGGCTTCTGATTACATGAAGAAATCAGTTTCAAGACTATTTAGATCTC 1059
 QY 1032 TACACAGAAATGGAAGATTTCTTGAAGAGGGGAGGAACAGATCTATTTAACTCGGG 1091
 Db 1060 TACAAGGATTATGAAGAAAGAAATTTGCTAGTGGCGGAGATCTCATATTTCTGCCATGCA 1119

QY 1092 AAAGATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAATGGCAATAG 1151
 Db 1120 ATAGAAGAATGAAGAAGTAGTAAGAATAATTAATGTCGAGTCAACATGTTTATTGAA 1179
 QY 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGGGTCTAAC 1211
 Db 1180 GGATATACGCCACCTGTTTCTGATACCTAAGCAATGCATAGCACTACACATATTAC 1239
 QY 1212 CTGCTTACAACTGTTTATCTTTGTCATGATGATATTTCAAAAAGAGTCTGTCGAA 1271
 Db 1240 TACCTCGGACAACTGCTATTTTGGGCATGAAGTCTGC---CACGGAGCAAGATTTTGAG 1296
 QY 1272 TGGGCTGCTCTGCACTCTCTTTTATAGATCTCAGGTATCTTGGTTCGAGCCCTAAT 1331
 Db 1297 TGGTTGTCAAAGAAATCCAAAATTTCTTGAAGCTAGTGTAAATTAATGTCGAGTTATCGAT 1356
 QY 1332 GATCTCATGACCCACCAAGCGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCTTCAAAAGT 1391
 Db 1357 GACACAGCCAGCTACGAGGTTGAGAAAGCAGGGGACAAATTTGCACTGGAATTCAGTGC 1416
 QY 1392 TATATGAAGAAATATAATGTCAATGAGGATATGCCCAACCTTGATTTTCAAGGAAGTA 1451
 Db 1417 TGCATGAGATTTATGTTATATCAACAAAGAGGCAATGGCTAAATTTCAAAATATGGCT 1476
 QY 1452 GAAGTGTGTGAAGATATAAACCAGAGATACCTCAACTAAAAACATTTCCAAAGCGG 1511
 Db 1477 GAGACAGCATGGAAGATATTAATGAAGGACTTCTTAGGCCCATCTCCGCTCTCTACAGAA 1536
 QY 1512 TTATTCATGCTGCTGATCTATTTGTGCCAGTTTCTTGAAGTTCAATATGCAAGAAA---G 1568
 Db 1537 TTTTAACTCTTATCTCATCTTGTCTGATTTGTTGAGGTTACATATACACATCTA 1596
 QY 1569 GATACTTACAGTATGAGGAGCAAGATACAACATCTCAATAAGTCTCTACTCTGTTAT 1628
 Db 1597 GATGGATACACTCATCGGAGAAAGTCTTAAACCTCACATTTAATTAACCTACTTGTGGAC 1656
 QY 1629 CCTATGATATATGA 1643
 Db 1657 TCCATCAAAATTTGA 1671

RESULT 15

US-09-895-752-1
 ; Sequence 1, Application US/09895752
 ; Patent No. 6559297
 ; GENERAL INFORMATION:
 ; APPLICANT: Chappell, Joseph
 ; APPLICANT: No. 6559297, Joseph P.
 ; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASES
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/895,752
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/398,395
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/100,993
 ; PRIOR FILING DATE: 1998-09-18
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; PRIOR FILING DATE: 1999-08-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1671
 ; TYPE: DNA
 ; ORGANISM: Nicotiana tabacum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (25)...(1668)
 US-09-895-752-1

Query Match 19.4%; Score 320.6; DB 4; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.le-84;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

Qy 15 GAAGAAACCTATTGCGCCCAATTGCCAACTTCTCCCAAGCATTTGGGGAGATCAGTTT 74
Db |||||
Qy 52 GAAGAAAGATGTTGCGCCCGCTGCCCGACTTCCCTAGTTTGGGGGTGATCAGTTTC 111
Db |||||
Qy 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAAA 134
Db |||||
Qy 112 CTTTCATTCTCAATTA--AAATCAGGTTGAGAAAGATAGTCTCAAGAGATTTGAAGCA 168
Db |||||
Qy 135 GAAGTGGCGCACTACTAAAAGAACTTTGGATATTTCCCTATGAACCATCCCAATTTGTTG 194
Db |||||
Qy 169 TTGAAGGAACAAACGAGGAATATGTTGTAGCAACTGGAATGAAATGSCCTGATACACG 228
Db |||||
Qy 195 AAGCTGATGATGAATTAACAGCCTTGGAAATACCGTATACCTTTGAACGGGAGATGAT 254
Db |||||
Qy 229 AATTGTAGACACTAATTGAACGCTTGGCATATCTACCACTTTGAGAAGAAATTTGAT 288
Db |||||
Qy 255 CATGCATTGCAATGATTTATGAACATATGTTGATTAACCTGGAATGGTGCCGCTCTCC 314
Db |||||
Qy 289 GATATTTGGATCAGATTTAACCAAACTCAAACTGCAACGATTTGTGCACTTCTGCA 348
Db |||||
Qy 315 TTATGGTTCCTTTATGCGAAGCAAGGATATTTATGTTACATGATGATGTTTTCATAAAC 374
Db |||||
Qy 349 CTTCAATTTGATGCTCAGGCAACATGCTTTCAACATCTCTCCCTGAATTTTCAGCAA 408
Db |||||
Qy 375 TATAAGCAAAATAGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCTT 434
Db |||||
Qy 409 TTCCAAAGCAAAATGGCAAAATTCAGGAATCTTGTCTAGTGAATGCTTTAGGATTTATG 468
Db |||||
Qy 435 GAGTTGTACGAAGCAACTCTATGAGGATCTCTGCGGAGATATATAGAAATGCTCTT 494
Db |||||
Qy 469 AACTTGTATGAAGCTTCAATGTAAGGACTCATGCTGAGATATCTTAGAAGCGCACTT 528
Db |||||
Qy 495 GGTTTTACACGATCTCGCTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCCGCT 554
Db |||||
Qy 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCACTCCACATTTGAAATCTCCA 579
Db |||||
Qy 555 CTTTTCACGAAATACACGGGCACTAAGCAACCCCTTTGGAAAGGTTGCGCAAGATA 614
Db |||||
Qy 580 CTTAGGAGCAAGTGACACATGCCCCCTTGAGCAATGTTTGCACAAGGGTGTCTCAGATC 639
Db |||||
Qy 615 GAGCGGCGCAGTACATTCCTT---TCTATCAACAAGATTCATCAACAAGACTTTA 671
Db |||||
Qy 640 GAGACCGGATCTCTCATCTCATCAATCATGACAGGAACATCGAAGATTAATGTGTTA 699
Db |||||
Qy 672 CTTAAACTTGCTAAGTTAGATTCATTTGCTTCAGTCAATGCAACAAGGAGAGCTCAGC 731
Db |||||
Qy 700 CTTGATTTGCCAAATTTGATTTCACTTGTCCAGATGTTGCAACAACAAGACTTGTCT 759
Db |||||
Qy 732 CATGTTGCAAAATGTTGGAAGCTTTCGATATCAAGAAAGACGCACTTGTGTTAGAGAT 791
Db |||||
Qy 760 CAAGTATCAAGGTGGTGGAAAGATTTGGATTTTGTAAACAACACTTCCATATGCTAGAT 819
Db |||||
Qy 792 AGAATGTTGAATGCTACTTTTGGGAGCTTAGTTCAGGCTATGAGCCACAGATTTCCCGG 851
Db |||||
Qy 820 CGAGTAGTTGAATGCTACTTTTGGGCATTTAGAGTTTATTTGAGCCTCAATFACTCTCA 879
Db |||||
Qy 852 GCTAGAGTTTCTTCAAAAAGCTTGTGTTGTTATAAATCTTTATAGATGACACTTATGAT 911
Db |||||
Qy 880 GCTCGGTCATGCTCGTTTAAGACCATATCAATGATTTGATTTGATGATGACACCTTTGAT 939
Db |||||
Qy 912 CCGTATGTTGATCTATGAAGACTTAAGATCTTTACTGAAGCTGTTGAAAGGTTGGTCAAT 971
Db |||||
Qy 940 GCTTACGGTACAGTTAAAGAACTTGGGCATACACAGATGCCATACAAAGATGGGATATC 999
Db |||||
Qy 972 ACATGCTTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA 1031
Db |||||
Qy 1000 AACGAAATTTGATCGGCTTCTGATTTACATGAAATACGTTTCAAAAGCTATTTAGATCTC 1059
Db |||||
Qy 1032 TACACGAAATGGAGAAATTTCTTGCAAGGAGGGAAGAACAGATCTATTTAACTCGCGC 1091
Db |||||

Db 1060 TACAAGGATTTATGAAGAAGATTTGCTTAGTGCCGAGAGATCTCATATTTGCTGCCATGCA 1119
Qy |||||
Qy 1092 AAGAATTTGTGAAAGAGTTTGTAGAAACCTGTAGTTGTTGAAGCAAAATGGGCAATAGAG 1151
Db |||||
Db 1120 ATAGAAAGATGAAGAAAGTAGTAAGAAATTTATAATGTGAGTCAACATGTGTTTATGAA 1179
Qy |||||
Qy 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAAATCATTTACTGGGGTGTCTAC 1211
Db |||||
Db 1180 GGATATACGCCACTGTTTCTGAAATACCTAAGCAATGCACTAGCAACTACCAATATTAC 1239
Qy |||||
Qy 1212 CTCCTTACAACTTTGTTATCTTGGCATGAGTGATATATTCAAAAAAGAGTCTCTCGAA 1271
Db |||||
Db 1240 TACCTCGGACAAACATCGTATTTGGGCATGAAGTCTGC--CACGGAGCAAGATTTTGAG 1296
Qy |||||
Qy 1272 TGGGCTGTCTCTGCACCTCTCTTTTAGATACCTAGTACTTGTGCGAGCCCTAAAT 1331
Db |||||
Db 1297 TGGTGTCAAGAAATCCAAAAATTTCTTGAAGCTAGTGTAAATTTATATGTCGATTCGAT 1356
Qy |||||
Qy 1332 GATCTCATGACCCACAGGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCTTCAAAGT 1391
Db |||||
Db 1357 GACACAGCCACGTACGAGGTTGAGAAAGCAGGGGACAAATTGCAACTGGAAATGAGTGC 1416
Qy |||||
Qy 1392 TATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAACTTTGATTTACAAGGAAGTA 1451
Db |||||
Db 1417 TGCATGAGAGATTATGGTATATCAACAAAGAGGCAATGGCTAAATTTCAAATATATGGCT 1476
Qy |||||
Qy 1452 GAAGATGTGGAAGAAAGATATAAACCGAGAGTACCTCACTCACTAAAAACATTTCCAGGCGG 1511
Db |||||
Db 1477 GAGACAGCATGGAAGAGATATTAATGAAGGACTTCTTAGGCCCACTCCCGTCTCTACAGAA 1536
Qy |||||
Qy 1512 TTATTGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAAGTTTCAATATGCGAGGAAA---G 1568
Db |||||
Db 1537 TTTTAACTCCTAATTTCTCAATCTTGTCTGTTTGTGGGTTACATATATACAAATCTA 1596
Qy |||||
Qy 1569 GATAACTTCACTGATGGAGAGCAATACAAAACATCTCAATAAGTCTCTACTCGTTTAT 1628
Db |||||
Db 1597 GATGGATACACTCATCCGGAGAAAGTCTTAAACCTCACATTTAATTAACCTACTTGTGGAC 1656
Qy |||||
Qy 1629 CCTATGAGTATATGA 1643
Db |||||
Db 1657 TCCATCAAAATTTGA 1671

Search completed: June 7, 2004, 12:41:20
Job time : 159 secs

	Query Match	61.7%;	Score 1017;	DB 16;	Length 1641;
	Best Local Similarity	76.2%;	Pred. No. 2.2e-263;		
	Matches 1251;	Conservative	0;	Mismatches 390;	Indels 0; Gaps 0;
QY	3	ATGGCACTTACAGAGAAACCTATTGCGCCCATTCGCAACTTTCTCTCCAGCATTTGG	62		
DbB	1	ATGCGCCTGACCGAAGAGAAACCGATCGCCCGATCGCTAACTTCCGCGCGTCTACTCG	60		
QY	63	GGAGATCATGTTTCTCATCTCTACAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGAAT	122		
DbB	61	GGTGACCAAGTTCCTGATCTTAGAAAGCAGGTTGAGCGGGTGTGAACAGATCGTAAAC	120		
QY	123	GATTTAAAAAAGAGTGGGCACTACTCTAAAAAGAAAGCTTTTGGATATTCCTATGAAACAT	182		

No.	Score	Match	Length	DB	ID	Description
1	1017	61.7	1641	16	US-10-411-066-37	Sequence 37, Appl
2	509.6	30.9	1644	16	US-10-411-066-39	Sequence 39, Appl
3	357.6	21.7	2024	9	US-09-887-886A-51	Sequence 51, Appl
4	357.6	21.7	2024	9	US-09-903-012-51	Sequence 51, Appl
5	357.6	21.7	2024	13	US-09-900-797-51	Sequence 51, Appl
6	357.6	21.7	2024	13	US-09-893-820-51	Sequence 51, Appl
7	351.6	21.3	1937	13	US-10-424-599-43432	Sequence 43432, A
8	330.4	20.0	1944	9	US-09-887-886A-31	Sequence 31, Appl
9	330.4	20.0	1944	9	US-09-903-012-31	Sequence 31, Appl
10	330.4	20.0	1944	10	US-09-900-797-31	Sequence 31, Appl
11	330.4	20.0	1944	13	US-09-893-820-31	Sequence 31, Appl
12	320.6	19.4	1671	9	US-09-887-886A-1	Sequence 1, Appl
13	320.6	19.4	1671	9	US-09-903-012-1	Sequence 1, Appl
14	320.6	19.4	1671	10	US-09-900-797-1	Sequence 1, Appl

Db 121 GACCTGAAGAAGAGTTCTGTCAGCTGCTGAAAGAGCTCTGGACATCCCGATGAACAC 180
Qy 183 GCCAATTTGTTGAAGCTGATGATGAATTTCAACGCTTTGGATACCGTATACCTTTGAA 242
Db 181 GCTAACTCTGCTGAACCTGATGACGAGATCCAGCGTCTGGGTATCCCGTACCACCTTCGAA 240
Qy 243 CGGAGATTGATCATGCTGCAATGTTATTTATGAAACATATGTTGATGAACCTGGAAATGTT 302
Db 241 CGCGAATCGACACGCTGCTGAGTGCATCTACGAACTTACGCGGACGACCTGGAACGCG 300
Qy 303 GACCGCTCTTCTTATGTTTCGCTTATGGAAGCAAGGATATATGTTTACATGTTGAT 362
Db 301 GACCGTCTTCTCTGTTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 363 GTTTTCAATTAATAAGCAAAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGTT 422
Db 361 GTTTTCAACTTACAGCAAGAACGCTGCTTCAACACGCTCTCTGGCTAACGACGTT 420
Qy 423 GAAGGTTTGTGTTGATGTTTACGAACTTCTATGAGGTTACCTGGGAGATTTATTA 482
Db 421 GAAGGCTGCTGGAACCTGACGAGGACCTCCATGCGGTGACCGGTGAATCATCTCTG 480
Qy 483 GAAGATGCTCTTGGTTTACAGATCTCTGCTTAGCATTTATGACAAAAGATGCTTTTCT 542
Db 481 GAGGCGGCTGGGTTTCAACCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 543 ARAACCCGCTCTTTTACCGAATAACAACGCGCACTAAAGCAACCCCTTTGGAAAGG 602
Db 541 ACTAACCCGCTCTGTTTCAACCGAATCAGCGTCTCTGAAACAGCGCTGTGGAAACGT 600
Qy 603 TTGCAAGATAGAGCGGCGGAGTACATCTCTTCTATCAACAACAAGATCTCATAC 662
Db 601 CTGCGCGTATCGAAGCAGCAGTACATCTCGTTTACCAGCAGGAGCTCTCACAC 660
Qy 663 AAGACTTTAACTTGAAGTTAGTTGATTTCAATTTGCTTTCAGTATGACAAAGGAA 722
Db 661 AAGACCTCTGCTGAACCTGCTGAAGTGAATTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 723 GAGCTCACCATGTTGCAATGTTGGAAGCTTTCGATATCAAGAAGACCGACCTTCT 782
Db 721 GAATGCTCTACGTTTGAAGTGGGAGGATTTGACATCAAGAAACCGCGCTGTC 780
Qy 783 TTAAGAGATAGAAATTTGTTGATGCTTACTTTTGGGACTAGTTTACGGCTATGAGCCACAG 842
Db 781 CTGCGTGACGATGCTGTTGATGTTTACTTCTGGGCTCTGGGTTCTGGTTATGACACAG 840
Qy 843 TATTCGCGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTATTAACCTTTATAGATGAC 902
Db 841 TACTCCCGTGCACGTTGTTCTTCACTAAAGCTGATGTTTATCACCCCTGATCGATGAC 900
Qy 903 ACTTATGATGCTGTTGATGTTTATGAGAACTTAAAGCTTTTACTGAAGCTGTTGAAAGG 962
Db 901 ACTTACGATGTTTACGCGACCTTACGAAAGAACTTGAAGATCTTTACTGAAGCTGTAGAACGC 960
Qy 963 TGGTCAATATGCTTGAACACACTTCCAGAAATACATGAACCCGATATACAAATTTATTC 1022
Db 961 TGGTCTATCACTTGTGCTGACACTCTGCGGAGTATGATGAACCCGATCTACAACTGTTTC 1020
Qy 1023 ATGGATATACACAGAAATGAAAGATTTCTTTGCAAGAGGAGGAGAAACAGATCTATTT 1082
Db 1021 ATGGATACCTACACCGAAATGAGGAAATCTCTGGCAAAAGAGGCGGTACCGACCTGTTTC 1080
Qy 1083 AACTGCGCAAGATTTGTTGAAGAGTTTGTAGAACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1142
Db 1081 AACTGCGGTAAGAGTTTGTAAAGATTTCTGTAAGTAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1140
Qy 1143 GCAATATGAGGACACATACCAACCTATGAAGAGCATGATCCAGTTGTTGTTGTTGTTGTTGTTGTTGTT 1202
Db 1141 GCTACGAGGCGCATATCCGACTACCGAAGACATGACCCCGTGTGTTATCATCACCGGC 1200
Qy 1203 GGTGTAACTGCTTACAACTATGTTTATCTTGGCATGAGTATATATCAAAAGAG 1262

Db 1201 GGTGCAAACTGCTGACCACTTGTCTATCTGGGTATGTCGACATCTTTTACCAAGGAA 1260
Qy 1263 TCTGTGCAATGGGCTGCTCTGACACCTCTCTTTTAGATACTCAGGTATATCTTGGTCTGA 1322
Db 1261 TCTGTTGAATGGGCTGTTTCTGCAACCCGCTGTTCCGTTACTCGGTATTTCTGGGTCT 1320
Qy 1323 CGCTAAATGATCTCATGACCCCAAGGCGGAGCAAGAAAGAAACAATAGTTATCTGAGC 1382
Db 1321 CGTCTGAACGACCTGATGACCCCAAGCAGAGCAGGAAACGTAACACACTCTTCTCTCTCT 1380
Qy 1383 CTGGAAGTTATGAGGAATATATGTCATGAGGAGTATGCCAAACCTTGAATTTTAC 1442
Db 1381 CTGGAATCTCTACATGAAGGAATATAACGTTAAACGAGAGTACGACACGACTCTGATCTAT 1440
Qy 1443 AAGGAAGTAAAGATGTTGCTGAAAGATATAAACCAGAGAGTACCTCAACAACCTAAAAACAT 1502
Db 1441 AAGGAAGTTGAAGCAGTATGGAAGACATCAACGCTGAATCTGACTACTTAAAAACATC 1500
Qy 1503 CCAAGCCGCTTATGATGCTGATCTATTTGTCAGTTTCTTGAAGTTCAATATGCA 1562
Db 1501 CCGCGCCGCTGCTGATGCGAGTAATCTACCTGTGCCAGTTCTCTGGAAGTACAGTACGCT 1560
Qy 1563 GGAAGGATTAACCTTCAACGTTATGGGAGACGATACAAACATCTCATAAAGTCTCTACTC 1622
Db 1561 GGTAAAGATAACTTCACTCGCATGGGCGGAGATACAAACACCTGATCAAAATCCCTGCTG 1620
Qy 1623 GTTTATCTCTATGATATATGA 1643
Db 1621 GTTTACCGATGTCATCTGA 1641

RESULT 2
US-10-411-066-39
; Sequence 39, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-000720
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Artemisia annua
US-10-411-066-39

Query Match 30.9%; Score 509.6; DB 16; Length 1644;
Best Local Similarity 57.3%; Pred. No. 1.8e-126;
Matches 942; Conservative 0; Mismatches 699; Indels 3; Gaps 1;

Qy 3 ATGGCACTTACAGAAACCTTATTCGCCCACTTCCCACTTCTCCAGCATTTGG 62
Db 1 ATGAGCCTGATCTAGAGGACGTAATCCGCCGAAACGCTAACTTCCCGTCTGAGATCTGG 60
Qy 63 GGAGATCAGTTTCTCATCTCTCAAAAGCAAGTAGCAAGGGTGGAAACAGATAGTGAAT 122
Db 61 GGTGACCACTTCTGGCATACGACGAGACGAGAGGTTGTAACAGGTTATCAGA 120
Qy 123 GATTTAAAAAGAGTGGCGCAACTACTAAAGAAAGCTTTGGATATTTCTTATGAAACAT 182
Db 121 GACCTGAAGAAAGAGTTAAATCTGAGCTGCTGACCGCTGTGAACCTCCCGACTCAGCAC 180
Qy 183 GCAATTTGTTGAGCTGATGTAATAATCAACGCTTGAATACCGTATCACTTTGAA 242

Qy	1323	CGCCTAAATGATCTCATGACCCCAAGCCGAGCAGAAAGAAACATAGTTTCATCGAGC	1382
Db	1321	ACTATGGACGACCTGGGCTCCCAAGAGCGAGGACCGTAAACACGTTGCTTCACT	1380
Qy	1383	CTTGAAAGTTTATGAAGGAATATAATGCAATGAGGAGTATGCCAAACCTTGATTAC	1442
Db	1381	ATCGAATGCTACATGAAGCAGTTTCAGCTAGCGAGCAGGACATACGAATCTCTGAAC	1440
Qy	1443	AAGGAAGTAGAAGATGTGTGAAAGATATAAACCGAGAGTAC---CTCACACTTAAAC	1499
Db	1441	AAGAAAGTTGAAGACGCGATGGAAAGAAATCAACCGTGAATCATGATCACTTGTAAAGAC	1500
Qy	1500	ATTCAGGCCGTTATTGATGGCTGTGATCTATTGTGCGAGTTTCTTGAAGTTCAATAT	1559
Db	1501	GTAACATCCATGTAGCGATGGGTACTGAACTTCTCCGCTCCGTTGACGTA CTGAC	1560
Qy	1560	GCAGGAAGAGTAATCTTCCACAGTATGGAGACGAATACAAACATCTCATTAAGTCTCTA	1619
Db	1561	AAGAACAAAGATCACTTCACCTCACCTTGCTGTAGAAGTTATCAACCACATCAAAATCCCTG	1620
Qy	1620	CTCGTTTATCCTATGAGTATATGA	1643
Db	1621	TTCGTTGACCGCATCATCACCTGA	1644

```

1  RESUOI 3
2  US-09-887-586A-51
3  ; Sequence 51, Application US/09887586A
4  ; Patent No. US20020094556A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Chappell, Joseph
7  ; APPLICANT: No. US20020094556A1, Joseph P.
8  ; APPLICANT: Starsks, Courtney M.
9  ; APPLICANT: Manna, Kathleen R.
10 ; TITLE OF INVENTION: SYNTHASES
11 ; FILE REFERENCE: 07678-025001
12 ; CURRENT APPLICATION NUMBER: US/09/887,586A
13 ; CURRENT FILING DATE: 2001-06-22
14 ; PRIOR APPLICATION NUMBER: 09/398,395
15 ; PRIOR FILING DATE: 1998-09-17
16 ; PRIOR APPLICATION NUMBER: 60/130,628
17 ; PRIOR FILING DATE: 1998-04-22
18 ; PRIOR APPLICATION NUMBER: 60/150,262
19 ; PRIOR FILING DATE: 1998-08-23
20 ; NUMBER OF SEQ ID NOS: 58
21 ; SOFTWARE: FastSeq for Windows Version 3.0
22 ; SEQ ID NO 51
23 ; LENGTH: 2024
24 ; TYPE: DNA
25 ; ORGANISM: Lycopersicon esculentum
26 ; FEATURE:
27 ; NAME/KEY: CDS
28 ; LOCATION: (32)...(1675)
29 ; OTHER INFORMATION: VFVT germacrene C synthase
30 ; US-09-887-586A-51

```

	Query Match	21.7%;	Score 357.6;	DB 9;	Length 2024;
	Best Local Similarity	52.7%;	Pred. No. 2.1e-85;		
	Matches	861;	Mismatches 0;	Indels 24;	Gaps 3;
	Conservative				
QY	29	TCGCCCAATTGCCAACTTTCCTCCAAAGCAATTTGGGGAGATTCAGTTTCTCATCTATCAAAA	88		
DB	58	TCGCCCCCTGGCGTAAATTTTCAACCATCTCTTTGGGGATATCATTTCTCTTTATATCTCA	117		
QY	89	GCAAGTAGACGCAAGGGTGGAAACAGATAGTAATGATTTAAAAAAGAAAGTCGGCGCAACT	148		
DB	118	TGAAATTACTAATCAAGAAAAAGTTGCAAGTTGATGAGTACAAAGAGACAAATTAGAAAAAT	177		
QY	149	ACTAAAGAAGCTTTGGATATTCCTATATGAACAATATGCCAAATTTGTTGAAGCTGATTGATGA	208		
DB	178	GCTGTTGGAACCTTCGCAAAATAGACACTCAAAA-----GCTTGTGTGTATAGACGC	228		

209 AATTCAACGCCCTTGGAAATACCGTATCACCTTTGAACGGGAGATTGATCATGCTATGCAATG 268
229 GATGCAACGATTGGAGTGGCTTATCATTTGATATGAAATTTGAAACATCCATTCAAAA 288
269 TATTATGAACAT ----- ATGGTGATAACTGGAATGGTGAACGCTCTTCCTT 316
289 CATTTTGTATGATCGTCCAAACAGAAATGATAATGACAAACACCTTTACGTTGTCTCT 348
317 ATGGTTCGCTTATGCGAAACGAGGATATTATGTTATCATGTGATGTTTTCATTAACCTA 376
349 TCGTTTTCGACTTGTGAGCAACAAGGCCATTACATGTTCTTCAGATGTGTGTCGAAGCAAT 408
377 TAAAGACAAAATGAGCGCTTCAAGCAATCGTTAGCTAATGATGTTGGAAGGTTTGTCTGA 436
409 CACCAACCAAGATGGAATTTCAAGGAACACTTACTAATGATGTCGAAGATATTGAG 468
437 GTTGTACGAAGCAACTTCTATGAGGCTACCTGGGAGATTTATTAAGAAGATCTCTTGG 496
469 TTTGTATGAAGCATCACATCTGAGAGTGGTAAATGAGGAGATTTCTTGAAGAGCTCTTAC 528
497 TTTTACAGATCTGCTTGTAGCATTTATGACAAAGATGCTTTTCTACAAACCCGCTCT 556
529 ATTTACCACTCATCTCGAGCTTATGCTTCTCAACTTGAAGCAATTAATAAATCTCT 588
557 TTTTACCGAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAAGATAGA 616
589 TAAGTTGAGTTGGTGAAGCTTAACTCAGCTATTGCACTATTACCAAGATGGG 648
617 GCGCGCGAGTACATCTCTTTCTATCAACCAAGATTTCTATPAACAAGACTTTTACTTAA 676
649 AGCTAGAAAATACATATCCATTTACGAAAACATGATGACACACCACTTTGCTTTTGA 708
677 ACTTGTAAATGAGTTCAATTTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 736
709 ATTTGCTAAATGGATTTTAACTGCTGCAAAAGTTTCAACAAAGAGCTTTAGTATCT 768
737 GTGCAATCGTGMAAGCTTTCATATCAAGAAAGCGCACTTTTAAAGATAGAAT 796
769 TACAAGTTGGTGAAGATTTGGAATTTGCAATAAATATCCATATGCAAGACAGGTT 828
797 TGTGATGCTACTTTTGGGAGTAGGTTAGGCTATGAGCCACAGATTTCCCGGGCTAG 856
829 GGTGAGTGTACTTCTGATATTTAGAGTGTATTTGAGCCAAATATATAGTGGTGGAG 888
857 AGTTTCTTCAAAAAGCTTGTCTGTTTAACTCTTATAGTATGACACTTATGATGCTA 916
889 AAAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTATGCTTA 948
917 TGTGATTTATGAAGAACTTAAAGATCTTACTGAAGCTTTGAAAGGTGGTCAATTTACATG 976
949 TGCACCTTTGACAACTTGTGACTTTCAATGATGCAATCCAGAGATGGGATGCTAATGC 1008
977 GTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATGATGATACATAC 1036
1009 AATTGATTCAATACAAACCATATATGACCTGCTTATCAAGCTTTCTAGACATTTACAG 1068
1037 AGAATGGAAGATTTCTTCAAGGAGGGAAGAACAGATCTATTAACTGGCGCAAGA 1096
1069 TGAATGGAACAAAGTGTGTCGAAGAGGTAACTGGACCGGTATATCTATATGCAAAAA 1128
1097 ATTTGGAAGAGCTTTGTTAGAAACCTGATGTTGGAAGCAAAATGGCAATGA ---GGG 1153
1129 TGAGATGAAAAGTTGGTGAAGGCTTATTTAAGGAACCCCAATGGTTGATGATTGGA 1188
1154 ACACATACCAACCACTGAAGAGCATGATCCAGTTGTTAATCATTAATCTGGCGGTGTACCT 1213
1189 CCATATTTCCAAAATATGAGCAACAAGTGGAGAAATGCAATCGTAAGTGTGGCTATATGAT 1248
1214 GCTTACAAACACTGTTTATCTTGGCATGAGTATATTTACAAAAGAGTGTGCGAATG 1273
1249 GATATCAACAACTTGTGTCGCTATGAGAAATTTATATCCACGAGACTTTTGAATG 1308
1274 GGCTGTCTGCGACCTCTCTTTTATGATCTCAGGTATATCTTGTGTCGACGCTTAAATGA 1333

1309 GTTGTGATGAATGAGTCTGTGATTGTTTCGAGCTTCCGCAATTTGTCGAGCAATGAACGA 1368
1334 TCTCATGACCCACCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCTTTGAAAGTTA 1393
1369 TATTTGTTGGACATGACATGAACCAAGAAAGAGGACATAGCTTCACTTATTGAATGTTA 1428
1394 TATGAAGGAATAAATCTCAATGAGGAGTATGCCCAACCTTGTATTACAAAGAAAGTAGA 1453
1429 CATGAAGATTTATGAGCTTCAAGCAAGAGACITACATTAAGTTTCTGAAAGAGGTAC 1488
1454 AGATGTGTGAAGATATAAACCGAGAGTACTCTCAACACTTAAACATTTCCAAAGCCGTT 1513
1489 CAATGCTATGGAAGGACATAAACAACAATCTCCCGTCCAACTGAAGTACCAATGTTGT 1548
1514 ATTTGATGGCTGTGATCTTATTTGTGCGAGTTTCTTGAAGTTCAATATGCGAGAAAGATAA 1573
1549 CCTTGAACGAGTTCTAAATTTTGACACGTTGGCTGACACGTTTATATAGGAGAAAGATAC 1608
1574 CTTTCAACGCTATGGGAGACGAATACAAACATCTCATAAAGTCTCTTACTCGTTTATCTAT 1633
1609 ATATTCAACGCCCAAGGAAACCTTTAAACCATGATTAAATCAATTAATTTGAATCTGT 1668
1634 GAGTATATGAGGAT 1647
1669 CAAAATATAAATAT 1682

RESULT 4
US-09-903-012-51
Sequence 51, Application US/09903012
Patent No. US20020094557A1
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. US20020094557A1, Joseph P.
APPLICANT: Starks, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTHASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/903,012
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/100,993
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 51
LENGTH: 2024
TYPE: DNA
ORGANISM: Lycopersicon esculentum
FEATURE:
NAME/KEY: CDS
LOCATION: (32)...(1675)
OTHER INFORMATION: VFNT germacrene C synthase
US-09-903-012-51

Query Match 21.7%; Score 357.6; DB 9; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
QY 29 TCGCCCAATTCGCAACTTCTCCAGCATTTGGGAGATTCAGTTTCTCATCTATCAAAA 88
DB 58 TCGCCCTTGGCTAATTTTCCACCATCTGTTGGGATATCATTTCTTTTATACTCA 117
QY 89 GCACTAGAGCAAGGGTGGAGACAGATAGTGAATTTTAAAGAAAGAGTGGCGCAACT 148
DB 118 TGAATTAATAATCAAGAAAAGTTGAGTTGATGATCAAGACACATTTAGAAAAT 177

149 ACTAAGAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTGAAGCTGATTGATGA 208
178 GCTGGTGAAGAACTTTCGCAATATAGCACTCAAAA-----GCTTGTGTTGATAGACG 228
209 AATTCAACGCTTTGGAATACCGTATCACTTTGAAACGGAGATTGATCATGCAATGCAATG 268
229 GATCAACGATTGGGAGTGGCTTATCAITTCGATTAATGAATGAATGAACATCCATTCAAAA 288
269 TATTATGAACAT-----ATGGTGAATGCTGGAATGGTGAACCGCTCTTCCTT 316
289 CATTTTGTATGCATCGTCCAAACAGATGAATGACAAACCACTTTACGTTGTGTCTCT 348
317 ATGTTCCGCTTATGCGAAGCAAGGATATTTATGATCATGTGTTTCAATTAACATTA 376
349 TCGTTTTCGACTTGTGAGGCAACAGGCGCATTTACATGCTTCAGATGTTTCAAGCAAT 408
377 TAAAGACAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCTTGA 436
409 CACCAACCAAGATGGAAATTCAGGAACACTTACTAATGATGATCCCAAGGATATTGAG 468
437 GTTGTAGCAACAACTTCTATGAGGGTACCTGGGAGATTAATTAGAAAGATGCTCTTGG 496
469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATTCCTTGAAGAGCTCTTAC 528
497 TTTTACAGCATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCGCTCT 556
529 ATTTACCAACCACTCATCTCGAGTCTATTGTCFCAACTTGAACAATAAATACTCTCT 588
557 TTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAAGAATAGA 616
589 TAAGTTGAAGTTGGTGAAGCCTTAACCTGAGCTATTGCGATGACTTTTCAAGGATGG 648
617 GCGCGGAGTACATTCCTTTCTATCAACAACAAGATCTCTATACAAAGCTTTTACTTAA 676
649 AGCTAGAAAATACATATCCATTTACGAAAACAATGATGCACACCACTTCTTTTGA 708
677 ACTTGCTAAGTTAGTTCAATTTGCTTCACTTTCAGTTCATGACAGGAGAGCTCAGCCAT 736
709 ATTTGCTAAATGGATTTTAAATGCTGCTCAAAAGTTTCCACCAAGAGAGCTTAGTGAT 768
737 GTGCAATGGTGGAAAGCTTTTCGATATCAAGAAAGACCGCCTGTTTAAAGAGATAGAT 796
769 TACAAGTGGTGGAAAGATTGGATTTGCAATTAATATCCATATGCAAGAGACAGTT 828
797 TGTGAATGCTACTTTTGGGAGTACCTGAGTTCAGGCTATGAGCACAGTATTCCTGGGCTAG 856
829 GGTGAGTGTACTTCTGATATTAGGAGTGTATTTGAGCCAAATATAGTCGTGCGAG 888
857 AGTTTCTTCAAAAAGCTGTTGCTGTTTAACTCTTATAGATGACACTTATGATGCGTA 916
889 AAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTGATGCTTA 948
917 TGGTACTTATCAAGAACTTAAGATCTTTTACTGAAGCTGTTGAAAGTGGTCAATTACATG 976
949 TGCACCTTTGACGAACTTTGACTTTCATGATGATGATCCAGATGGGATGCTAATGC 1008
977 OTTAGACACTTCCAGAAATACATGAACCGGATATACAAATTAATTCATGATACATACAC 1036
1009 AATTGATTCAATACAAACATATATGAGACCTGCTTATCAAGCTCTTCATGACATTTACAG 1068
1037 AGAAATGGAAGAAATTTCTTGAAGAGGAGGAGAAACAGATCTATTTAACTGCGGCGAAAGA 1096
1069 TGAATGGAAACAGTGTGTGTCGAAAGAGGTAAACTGGACCGGTGATATCTATGCAAAAAA 1128
1097 ATTTGTGAAAGAGTTGTTAGAAACCTGATGTTGAAGCAAAATGGGCAATGA---GGG 1153
1129 TGAGATGAAGAAAGTTGGTGAAGAGCTTATTTAAGGAACCCCAATGTTGCAATGATTGGA 1188
1154 ACACATACCAACCACTGAAGAGCATGATCCAGTGTGTAATCATTAATGCGGCGTCAACCT 1213
1189 CCATATTTCCAAAATATAGGAACCAAGTGGAGATGCAATCGTAAGTGTGCTGGCTATGAT 1248
1214 GCTTACAAACAACTGTTTATCTTGGCATGAGTGATATTTTCAAAAGAGTCTGTGCAATG 1273

Db 1249 GATATCAAAACATTTGCTTGGTCGGTATAGAGAAATTTATATCCACGAGACTTTTGAATG 1308
QY 1274 GGCTGTCTCTGCACTCTCTCTTTTATAGATATCTAGGTATATCTTGGTCGACGCTAAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTTGCGAGCTTCCGCAATGATGCGAGCAATGAACGA 1368
QY 1334 TCTCATGACCAACAGCCCGAGCAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAAGAAAGAGGACATGTAGCTTCACTTATTGAATGTTA 1428
QY 1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTGATTTCACAGGAAGTAGA 1453
Db 1429 CATGAAGATATTGAGCTTCAAGCAAGAGACTTACATTAAGTTCCTGGAAGAGTCAAC 1488
QY 1454 AGATGTGTGGAAGATATAAACCGAGAGTACCTCACAACTAAAAACATTCCAAGCCCGTT 1513
Db 1489 CAATGATGGAAGAGACATAAAACAACTTCTCCGTCCTCAACTGAAGTACCAATGTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTTTGTGCGAGTTTCTTGAAGTTCATATGCAAGGAAGATTA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGCTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
QY 1574 CTTTCAACGATGGGAGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACCGCAAGGAAACTTTAAACATGATTAATCCAATACTAATTGAATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAATAT 1682

RESULT 5

US-09-900-797-51
; Sequence 51, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germaecene C synthase
US-09-900-797-51

Query Match 21.7%; Score 357.6; DB 10; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCATTTGCCAATTTCTCCTCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCATTTGGCTTAATTTTCAACCATCTGTTGGGATATCATTTCTCTTATCTCA 117
QY 89 GCAAGTAGAGCAAGGGTGGCAACAGATAGTAATGATTTAAAAAAGAGTGGCGCAACT 148

Db 118 TGAATTACTAATCAAGAAAAGTTGAATTGATGAGTACAAAGAGACAATTTAGAAAAAT 177
Qy 149 ACTAAAAGAGCTTTGGATATTCCTATGAAACATGCCAAATTTGTTGAAGCTGATTCATGA 208
Db 178 GCTGGTGAACACTTGGCAATAGCACTCAAAA-----GCTTGTGTTGATAGACGC 228
Qy 209 AATTCAAGCCTTGGATACCGTATACCTTTGAACGGGAGATTGATCATGATTCGAATG 268
Db 229 GATCAACGATTTGGAGTGGCTTATCATTTCCGATTAATGAATTTGAACATCCATTCAAAA 288
Qy 269 TATTATGAACAT-----ATGGTGATAACTGGATGTGACCGCTCTTCCTT 316
Db 289 CATTTTGTATGATCGTCCAAACAGAAATGATATGACAAACCTTTAGCTGTGTCTCT 348
Qy 317 ATGTTCCGCTTATGCGAAAGCAAGGATATTTATGTTACATGTGATGTTTTCATTAACCTA 376
Db 349 TCGTTTTCGACTTGTGAGCAACAGGCCATTTACATGCTTCCAGATGCTTCAAGCAATT 408
Qy 377 TAAAGCAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTGCTTGA 436
Db 409 CACCAACCAAGATGGGAATTTCAAGAAACACTTACTAATGATGTCCTCAAGGATTTAG 468
Qy 437 GTTGTACGAACCACTTCTATGAGGTTACCTGGGAGATTTATATGAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCAATCTGAGAGTGGTAAATGAGAGATTTCTTGAAGAGCTCTTAC 528
Qy 497 TTTTACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCGCTCT 556
Db 529 ATTTACCACTCATCTCGAGTCTTGTCTCCAACTTGAGCAATAATAATACTCTCT 588
Qy 557 TTTTACGAAATACACGGGCACTAAAGAACCCCTTTGGAAAAGGTTGCCAAGAAATAGA 616
Db 589 TAAGGTTGAAGTTGGTGAAGCTTAACTCAGCCTATTTCGATGACTTTACCAAGGATGG 648
Qy 617 GCGCGCGAGTACATTCCTTCTATCAACAAGATTTCTCATACAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTAGAAAACATGATGACACACCCACCTATTCCTTTGAA 708
Qy 677 ACTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCATTGCAACAAGAGAGCTCAGCCATGT 736
Db 709 ATTTGCTAAATTTGATTTTAACTGCTGCAAAAGTTTCAACAAGAGAGCTTTAGTATCT 768
Qy 737 GTGCAATGGTGAAGCTTTGATATCAAGAAAGCGCACTGTTTGAAGAGATGAAAT 796
Db 769 TACAAGTGGTGAAGATTTGGATTTTGCATAATTAATATCCATATGCAAGAGACAGGTT 828
Qy 797 TGTGAATGCTACTTTTGGGAGTAGTTTCAAGCTATGAGCCACAGTATTCCTCGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATTAGAGTGTATTTTGGCCAAATATATAGTCGTCGAG 888
Qy 857 AGTTTCTTCAAAAAGCTTTGCTGTTTATAACTCTTATAGATGACACTTATGATCGTGA 916
Db 889 AAAAATGATGACAAAAGTACTCAACTGACCTCCATATTGACGACACTTTTGTATGCTTA 948
Qy 917 TGGTACTTATGAAGAACTTAAATCTTTACTGAAGCTGTTGAAGTGTGCTCAATACATG 976
Db 949 TGCACCTTTGACGAATTTGACTTTTCAATGATGCAATCCAGAGATGGGATGCTAATGC 1008
Qy 977 CTTAGACACATCCAGATACATGAACCGATATACAAATTTATTCATGATACATACAC 1036
Db 1009 AATTGATTCATACAAACATATAGAGCTGCTGTTATCAAGCTCTTCTAGACATTTTACAG 1068
Qy 1037 AGAAATGGAAGATTTCTTGAAGAGGAGGAACAGAGATCTATTTAACTGCGGCAAGA 1096
Db 1069 TGAATGGAACAAAGTTGTTCCAAAGAAAGTAACTGGACCGTGTATATATGCAAAAAA 1128
Qy 1097 ATTTGTGAAGATTTGTTAGAAAACCTGATGTTGGAACAAATGGGCAAAATGA---GGG 1153
Db 1129 TCAGATGAAAAGATTTGGTGAGAGCTTATTTTAAAGGAAACCCCAATGTTGAATGTTGTA 1188
Qy 1154 ACACATACCAACCACTGAGAGCATGATCCAGTTGTAATCATCTACCTGCGGTGCTAACCT 1213
Db 1189 CCATATTCAAAATATGAGAAACAAGTGGAGAAATCGTAAGTGTCTGGCTATATGAT 1248

Qy 1214 GCTTACAAACAACTTGTATCTTGGCATGAGTGATATATTCACAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACTTGTCTTGGTCGGTATAGAAAGATTTATATCCACGAGACTTTTGAATG 1308
Qy 1274 GCTGTCTCTGCACCTCTCTTTTATAGACTACTCAGGTATACTTGGTCGACGCTTAAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTTGAGCTTCGGCATTTGATTCGCCAGAGCAATGAACGA 1368
Qy 1334 TCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCATCGAGCCTTTGAAAGTTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAGAAAGAGACATGTAGCTTCACTTATTGAATGTTA 1428
Qy 1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTGTATTTACAAGAGATAGA 1453
Db 1429 CATGAAAGATTTATGGAGCTTCAAAAGCAAGAGACTTACATTAAGTTCTCTGAAAGAGGTAC 1488
Qy 1454 AGATGTGCGAAGATATAAACCAGAGTACCTCACAACATAAACAATTCACAAGCCGCTT 1513
Db 1489 CAATGATGAAGGACATAAACAACAATTTCTCCGTCCAACTGAAGTACCAATGTTTGT 1548
Qy 1514 ATTGATGGCTGTGATCTATTGTCGCCAGTTTCTTGAAGTTCAATATGACGAAAAGGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGTTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
Qy 1574 CTTCACACGTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACGCCCAAGGAAACTTAAAAACATGATTAATCCATACTAATTTGAATCTGT 1668
Qy 1634 GAGTATATGAGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 6
US-09-893-820-51
; Sequence 51, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Sparks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-893-820-51

Query Match 21.7%; Score 357.6; DB 13; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
Qy 29 TCGCCCCCAATGCCCACTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88

Db 58 TCGCCCTTGGCTAATTTTCCACCATCTGTTTGGGATATCATTTTCTTTCTTACTCA 117
Qy 89 GCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTTAAAAAAGAGAGTGGGCAACT 148
Db 118 TGAATTTACTAATCAAGAAAAGTTGAAGTTGATGATGATGATGATGATGATGATGATGAT 177
Qy 149 ACTAAAGAACCTTTGGATATCTCTATGAACATGCCAATTTGTTGAAGCTGATGATGA 208
Db 178 GCTGGTGGAAACTTTGCGCAATAGCACTCAAAA-----GCTTGTGTTGATGACGC 228
Qy 209 AATTCAACGCTTTGAAATACCGTATCACTTTGAAACGGGAGATTGATCATGCAATTCGAATG 268
Db 229 GATCGAAGATTGGAGTGGCTTATCATTTGATTAATGAAATGAAACATCATTCATAAA 288
Qy 269 TATTTATGAACAT-----ATGGTGAATGAACTGGAATGGTGAACCGCTCTTCTCT 316
Db 289 CATTTTGTATGATCGTCCAAACAGATGATGATGATGATGATGATGATGATGATGATGAT 348
Qy 317 ATGTTTCGCTTTATGCAAGACAGGATATTTATGTTATGATGATGATGATGATGATGAT 376
Db 349 TCGTTTTCGACTTTGAGGCAACAGGCACTTACATGTTCTCAGATGTTGTTCAAGCAATT 408
Qy 377 TAAAGACAAAATGGAGGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCCTTGA 436
Db 409 CACCAACCAAGATGGGAATTCAGGAACACTTACTAATGATGTTCCAAAGATTTATGAG 468
Qy 437 GTTGTACGAACCACTTCTATGAGGTTACCTGGGAGATTTATATGAGATGCTCTTGG 496
Db 469 TTTGTATGAACATCACTCTGAGAGTGGTAAATGAGGAGATTTCTTGAAGAGCTTTAC 528
Qy 497 TTTTACAGATCTGCTTTAGATTTATGACAAAAGATGCTTTTCTCAAAACCGCTCT 556
Db 529 ATTTTACCACCACTCATCTCGAGTCTATTGTTCTTCCAACTTTGACCAATAATAAATCTCT 588
Qy 557 TTTTACCGAATACACCGGCACTAAACCAACCCCTTTGGAAAAGTTGCGCAAGATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCCTTAACCTGAGCTATTGCGATGATTTTACCAGAGATGG 648
Qy 617 GCGGCGCAGTACATTTCTTCTATCAACCAAGATTTCTATACCAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTAGAAAACATGATGATGATGATGATGATGATGATGAT 708
Qy 677 ACTTGCTAAGTTAGATTTCAATTTGCTTCACTGATGATGATGATGATGATGATGATGAT 736
Db 709 ATTTGCTAAAATGGATTTTAACTGCTGCAAAAGTTTCAACCAAGAGAGCTTAGTATCT 768
Qy 737 GTGCAATGTTGGAAAGCTTTTCGATATCAAGAAAGCACTTTTAAAGAGATGAT 796
Db 769 TACAAGTTGGTGAAGATTTGGATTTTGGCAATTAATATCCATATGCAAGAGACAGGT 828
Qy 797 TGTGAATGCTTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGTATTTCCGGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATAGGAGTGTATTTGAGCCAAATATAGTCTGTGGAG 888
Qy 857 AGTTTCTTCAAAAGCTTTGCTGTTATTAATCTTTATAGATGACACTTATGATGGTA 916
Db 889 AAAAATGATGACAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTATGCTTA 948
Qy 917 TGGTACTTATGAGAACTTATGATCTTTTACTGAACTGTTGAAAGTGGTCAATTACATG 976
Db 949 TGCACCTTTGACGAACTTTGACTTTTCAATGATGCAATTCAGAGATGGATGCTAATGC 1008
Qy 977 CTTAGACACACTTCCAGAAATACATGAAACCGATATACAAATTTATTCATGATACATACAC 1036
Db 1009 AATTGATTTCAATACCAATATATGAGACCTGCTTATCAGCTCTTCTAGACATTTACAG 1068
Qy 1037 AGAATGGAAGAAATTTCTTGAAGAGGAGGAGAAACAGATCTATTATTAATGCGGCAAGA 1096
Db 1069 TGAATGGAACAAAGTGTGTGCAAAAGAGGTAACCTGGACCGTGTATATCTATGCAAAAA 1128
Qy 1097 ATTTGTGAAGATTTGTAGAACTGATGTTGAGCAAAATGGGCAATGA---GGG 1153
Db 1129 TGAGATGAAAAGTTGGTGAGAGCCTATTTTAAGGAAACCCCAATGGTTGAATGATTTGTA 1188

Qy 1154 ACATATACCAACCACTGAGAGCATGATCCAGTTGTATCTTGTATCTTACTGGCGGTGCTAACCT 1213
Db 1189 CCATATTCCAAATATGAGGAACAAGTGGAGAAATGCAATCGTAGTCTGGCTATATGAT 1248
Qy 1214 GCTTACAAACACTTTGTTATCTTGGCATGAGTATATTTCAAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACTTGTCTGCTGCTAGAGAAATTTATATCCACGAGACTTTTGAATG 1308
Qy 1274 GGCTGTCTGCACTCTCTCTTTTATAGTACTCAGGTATATCTTGGTCGACGCTAAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTTGTTCCGCTTCCGATGATTTGCCAGAGCAATGAACA 1368
Qy 1334 TCTCATGACCCACAAGCCGAGCAAGAAAGAAAACATAGTTTCATCGAGCCTTTGAAAGTTA 1393
Db 1369 TATTTGTGACATGAGATGAACAAGAAAGAGGACATGTAGCTTCACTTATTTGAATGTTA 1428
Qy 1394 TATGAAGATATTAATGTCAATGAGGAGTATGCCAAACCTTGTATTTACAAGGAAGTAGA 1453
Db 1429 CATGAAAGATTTATGGAAGCTTCAAAAGCAAGAGACTTACATTAAGTTTCCCTGAAAGAGT 1488
Qy 1454 AGATGTGTGAAAGATATAAACCAGAGATACCTCACAACATAAAACATTTCCAAAGCCGTT 1513
Db 1489 CAATGATGAGAGACATAAAACAATTTCTCCGTCCTCAACTGAAGTACCAATGTTGT 1548
Qy 1514 ATTTGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAAGTTCAATATGCGAGAAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGTTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
Qy 1574 CTTCACAGTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACCGCAAGGAAACCTTAAACACATGATTAATCCAAATACTAATTTGAATCTGT 1668
Qy 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAATAT 1682

RESULT 7
US-10-424-599-43432/c
; Sequence 43432, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43432
; LENGTH: 1937
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139218C.1
US-10-424-599-43432

Query Match 21.3%; Score 351.6; DB 13; Length 1937;
Best Local Similarity 53.0%; Pred. No. 8.6e-84;
Matches 865; Conservative 0; Mismatches 734; Indels 33; Gaps 4;

Qy 27 ATTGCGCCCATTTGCCAACTTTCTTCCAAAGCAATTTGGGAGATCAGTTTCTCATCTATCAA 86
Db 1815 ATTGCGAACTGCAAAATTTTCTCTTCCGTTTGGGCGGATTTTCTTTTACTATGTT 1756
Qy 87 AAGCAAGTAGACAAAGGCTGGAAACAGATAGTGAATGATTTTAAAAAAGAGAGTGGCGCA 146
Db 1755 CCTAGTTCCGTGGAAGATGATAGTACATCAACCAAGCTCAATCAACAAAGAGAGAGT 1696

QY 147 CTACTAAAGAGCTTTGGATATCTCTATGAACATGCCCAATTTGTTGAGCTGATTGAT 206
Db 1695 AGGAAGATGCTTATTGCTCTTATGATAACAAATTTCTATTATAATTAAGATTTATTGAC 1636
QY 207 GAAATTCACCGCTTTGGAATACCGTATCACTTTGAAACGGGAGATTGATCATGTCATTTGCCAA 266
Db 1635 TCAGTCCAAACGCTTGGGTGTTCTTACCACTTTTGAACATGAAATTTGATGGAGCAATTCAC 1576
QY 267 TGTATTTATGAACAATATGGTGATTAACCTGGAATGG-----TGAC 305
Db 1575 CAAATTTTCAACATTTCAACAAGGACCAATAATATCATAACTCACGATGATGCTTTGT 1516
QY 306 CGCTCTTCTTATGTTTCGCTCTTATGCGAAAGCAAGGATATATGTTTACATGATGATGTT 365
Db 1515 CATGGGCTTTACTCTTTTCGGTTCCTTAGGCAACAGGTTATCACTTTCAATCAAAATGTA 1456
QY 366 TTCAATAACTATAAAGACAAAATGAGCGTTCAAGCAATCGTTAGCTAATGATGTTTGAA 425
Db 1455 TTTTCAAAATTCAAAGACCAAACTAGAACTTTAGTGAAGAAAGCAGCCCAATGATATACAA 1396
QY 426 GGTTCGTTGAGTTGTAGCAAGCACTTCTATGAGGTTACCTGGGGAGATATATATTAGAA 485
Db 1395 GGAATGTTGAGCTTTGACGAAGCTGCTGAACCTAAGAAATGCATGGAGAGGATATACTTGAA 1336
QY 486 GATGCTCTTGGTTTACACGATCTCGTCTTAGCAATATGACAAAAGATGCTTTTCTACA 545
Db 1335 GAAGCACAAATTTTCGCTCTTCTGCTTCACTTAAGTCTCTAATACCCAAAT-----A 1282
QY 546 AACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTTGGAAAAGGTTG 605
Db 1281 AGTCCTTCTATGATTGCAACAAGTCAAGCATAGCTTAAAGACGATCACTTGGAAAAGGATTG 1222
QY 606 CCAAGAAATAGAGCGGGCGGAGTACATTCCTTTCTATFCAACAAACAGATTCCTCATACAAAG 665
Db 1221 CCTAGGCTGGAGGCCACATATATATATGCTTTCTACGAAGAAGATTCCTTCAATGATGAA 1162
QY 666 ACTTTACTTAACTGCTAAGTTAGAGTTCAATTTGCTTTCAGTCAATGTCACAGGAGAG 725
Db 1161 AAATGCTTAACTTTGCAAACTAGATTTTAAATGTTGCGAGAGTTATCATCAGAAGAA 1102
QY 726 CTCAGCCATGTTGCAAAATGGTGG---AAAGCTTTTCGATATCAAGAAGAACGCACTTTGT 782
Db 1101 GTTAAACATGTGACAGGTGGTGGATTAAGATTTAAATGTTCTCAACGAAATTCCTATTT 1042
QY 783 TTAAGAGATAGAAATTTGTAATGCTACTTTTGGGACTAGGTTTCAAGCTATGAGCCACAG 842
Db 1041 GTACAGATAGGATTCAGAGATGTTACTTTTGGATTTTGGGAATATATCTTTGAGCCACAA 982
QY 843 TATTCCCGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTATTAATCTTATAGATGAC 902
Db 981 TATTCTCTTGTAGAGGATAACAAACGAAAGTAATTTGCTCTATGTTCTGTCTATGATGAC 922
QY 903 ACTTATGATGGTATGTTACTTATGAAGAACTTAAAGATCTTTTACTGAAAGCTGTTGAAAGG 962
Db 921 ATGTATGATGCTATGGAACCACTTACGAACTTGGAGCTTTTCCCAATGCAATGGAGAGG 862
QY 963 TGGTCAATTAATGCTTTAGACACATTCAGAAATACATGAACCGGATATACAAATTTATTC 1022
Db 861 TGGGATATTTGTTGCTGGATGATCTCCAGAAATACATCAAAAGTATGTTATATAGAAAT 802
QY 1023 ATGGATACATACAGAAATGAAGAAATTTCTTGAAGAGGAGGAGACAGATCTATTT 1082
Db 801 TTGAATGTTTATGAAGAATAGAGGAAGATGAGAAAACAAGGAAAGTATATTGTCATC 742
QY 1083 AACTGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGTAAGCAAAATGG 1142
Db 741 AAGTATGCTAAGAAAGATGAAGATTAATCAAGGCTCACATGCTGAGGCAAGATGG 682
QY 1143 GCAATAGGGGACACATACCAACCTAGAGCAATGATCCAGTTGTGTAATCAATATCTGCG 1202
Db 681 CTTCAATTCATACACCGCTCAATAGAGGAGTACATGCAAGTTAGAAATGTATCAAGT 622
QY 1203 GGTGTAACCTGCTTACAACTCTGTTATCTTGGCATGATGATATATTCACAAAGAG 1262

Db 621 GGTTACTCTATGGTGATCACCATATGTTTGTGGCATGAAGATACAA---ACAGAGGAG 565
QY 1263 TCTGTGAAATGGGCTGTCTCTGACACCTCCTCTTTTGTAGATATCTCAGGTATATCTGGTGA 1322
Db 564 GTCCTTATATGGGCAACAAGTGTATCCCAATAATTTATGGGGCTGCTTCAATTTATTGTAGG 505
QY 1323 CGCCTAAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCTATCGAGC 1382
Db 504 CTTATGATGACATTTGTTGAAATGAGTTTGAGCAGGAAAGAACACGTTGTCATCAAGC 445
QY 1383 CTTGAAAGTTTATATGAAGGAATAATATGTCATGAGGAGTATGCCCAACCTTGTGATTAC 1442
Db 444 ATTGAAGCTTATATGAAGCAACATACACCTCAAGGCAAGATGCCATTAATAAATCTACT 385
QY 1443 AAGGAAGTGAAGATGTTGGAAGATATATAACCCGAGAGTACCTCACAACTAAAAACATT 1502
Db 384 GAGATGTTTAAAGAGTGTCTTGAAGGACATCAATGAGGATGCTTTAATCTCTACTGAAGTG 325
QY 1503 CCAAGCCGTTTATGATGGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCAATATGA 1562
Db 324 CCAATGAATTTCTTTTGGCTGTTGTCAACCTTGTGCGCATGATGACGTGCTTTTACAAA 265
QY 1563 GGAAGGATTAACCTTACACGATGATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTC 1622
Db 264 GATGAAGATTAATATACAAATGCAAGGAGGTTAATGAAGATTTACATCAAAACTTTTATTA 205
QY 1623 GTTTATCTCTATG 1634
Db 204 GTTAATAAGATG 193

RESULT 8
US-09-887-586A-31
; Sequence 31, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-887-586A-31

Query Match 20.0%; Score 330.4; DB 9; Length 1944;
Best Local Similarity 52.0%; Pred. No. 4.5e-78;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAACCTTATTCGCCCATTCGCAACTTCTTCCCAAGCATTTGGGGAGATCAGTT 73
Db 92 AGAGGAGGAGATGTTTCGCCCATAGCTACTTCTCTCCAAGTCTTTCGGGTGATCGTTT 151
QY 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAA 133

QY	1199	TGCGGTGCTAACTCTGCTTACACAACTTGTATCTTGGCATGAGTATATTCACAAA	1258
Db	1283	CAGCACTTATTACTTGTCTAGCACTACATCTTATTTTGGGCATGGAAGTCTCTAACAAAGCA	1342
QY	1259	AGAGTCTGTGCAATGGGCTGTCTCTGCACTCTCTTTTGTAGATACTCAGGTATACTTGG	1318
Db	1343	AGATTTTG---AATGGTTGGCAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG	1399
QY	1319	TCGAGCGCTAAATGATCTCATGACCCACAAAGCCGAGCAAGAAAGAAACATAGTTTCATC	1378
Db	1400	CCGAGTCATAGATGACATAGCACCTATGAGGTTGAGAAAGGGTAGAGGTGAGATTGCCAC	1459
QY	1379	GAGCCTTGAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTGTAT	1438
Db	1460	TGGAATTTGAATTTTACATGAGAGATTATGGTGTATCCACAGAAAGGCCATGGAAAAATT	1519
QY	1439	TTACAAGGAAGTAGAAGATGTGTGGAAGATATATAACCGAGAGTACCTCACAACTAAAAA	1498
Db	1520	CCAGAATAATGGCTGAGACAGCATGGAAGGATGTAATGAAGGAATCCTTCGACCACTCC	1579
QY	1499	CATTCCAAAGCGCGTTATTGATGGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCAATA	1558
Db	1580	CGTCTCTACAGAGATTCTCACTCGCATTTCTCAATCTTGTCTCGCATATCGATGTACTTA	1639
QY	1559	TGCAGGAA---AGGATACTTACACAGTATGGGAGACGAATACAAACATCTCATAAAGTC	1615
Db	1640	TAGACAATAACAGATGATACACTCATCCGGAAGAAAGTACTAAACCTCATATTATTGC	1699
QY	1616	TCCTACTCGTTTATCCTATGAGTATATGA	1643
Db	1700	GTTGTGGTGGACTCTATTGAAATTTAA	1727

RESULT 9

US-09-903-012-31

Sequence 31, Application US/09903012

Patent No. US20020094557A1

GENERAL INFORMATION:

APPLICANT: Chappell, Joseph

APPLICANT: No. US20020094557A1, Joseph P.

APPLICANT: Starks, Courtney M.

APPLICANT: Manna, Kathleen R.

TITLE OF INVENTION: SYNTHASES

FILE REFERENCE: 07678-025001

CURRENT APPLICATION NUMBER: US/09/903,012

CURRENT FILING DATE: 2001-07-11

PRIOR APPLICATION NUMBER: 09/398,395

PRIOR FILING DATE: 1999-09-17

PRIOR APPLICATION NUMBER: 60/100,993

PRIOR FILING DATE: 1998-09-18

PRIOR APPLICATION NUMBER: 60/130,628

PRIOR FILING DATE: 1999-04-22

PRIOR APPLICATION NUMBER: 60/150,262

PRIOR FILING DATE: 1999-08-23

NUMBER OF SEQ ID NOS: 58

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 31

LENGTH: 1944

TYPE: DNA

ORGANISM: Solanum tuberosum

FEATURE:

NAME/KEY: CDS

LOCATION: (57)...(1724)

OTHER INFORMATION: vetispiradiene synthase

US-09-903-012-31

Query Match 20.0%; Score 330.4; DB 9; Length 1944;

Best Local Similarity 52.0%; Pred. No. 4.5e-78;

Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAAAAGCTATTTCGCCCATTTCTCCAAAGCATTTGGGAGATCAATT 73

Db 92 AGAGGAGGAGATTGTTTGGCCCATAGCTGACTCTCTCCAAAGCTTTTGGGGTGATCGTTT 151
Qy 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGACAGATAGTAGTAATGATTTAAAAA 133
Db 152 TCATTCATTCCTCCCTCGCAATCAGATGCTGGAAAATATGCTCAAGAGATCGAAACTTT 211
Qy 134 AGAAGTGGGGCAACTACTAAAAAGAGCTTTGGATATTCCTATGAACATGCCAAATTTGTT 193
Db 212 GAAGGAACAATCAAGAAATATATATCTGCACTCTCTCGAAGAACATTTGGCTGAGAAAT 271
Qy 194 GAAGCTGATGATGAATAATCAACGCTTGGAAATACCGTATACCTTTGAAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGTTGAGCGCTTGGCATTTGCTTATCATTTTGAAGAAACAAATAGA 331
Qy 254 TCATGCATTCGAATGATTTATGAACATATGTT-----GATAACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTCAAGACAGATCTTAACCTTTGAGGCTCACGAGTACATGA 391
Qy 299 TGGTGACCGCTCTTCCTTATGTTTCCGCTCTTATGCGAAAGCAAGAGATATATGTTACATG 358
Db 392 TTTCACAACTTTATCGTTCAATTTGCACTTTGAGACAAACATGTTTACAATATCTCCCC 451
Qy 359 TGATGTTTCAATACTATAAGACAAAATGAGAGCGTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAAATTTTATTTAGATTCGAAGATCAAAAGGCAAAATTTAAAGAAATCTCTTTGTAACGA 511
Qy 419 TGTGAGAGGTTTGTGCTGTGACGAACAATCTTATGAGGAGTACCTGGGAGATTAT 478
Db 512 CATCAAGGCTCTTTGAACTTATCAAGGCTCGCATGTAAGGACTCATGGAAGAAATAT 571
Qy 479 ATTAGAAGATGCTCTTTGGTTTACAGATCTCTGCTTTAGCAATTTATGACAAAAGATGCTTT 538
Db 572 TTTGGAAGAGGCACATGCTTTCTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
Qy 539 TTCTACAAACCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCCTTGAGCAATCTCTCCATAA 682
Qy 599 RAGGTTGCCAAGATAGAGCGCGCAGTACATCTCTTCTATCAACAAGATTTCTCA 658
Db 683 GAGCAATCCAGAGTTGAGACAGCGTACTTCTATCTATCTAGAGAGGAGGAACAGAA 742
Qy 659 TAACAAGACTTTACTTTAAACTTCTTAAGTTAGAGTTCAATTTGCTTCAGTCAATGACAA 718
Db 743 GAATGATGTTGCTTCAATTTGCAAACTGAGTCACTTACTTACTTACAGATGTTGACAA 802
Qy 719 GGAAGAGCTCAGCCATGCTGTCMAATGTTGGGAAAGCTTTGATATCAAGAAAGCGACC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTTGGAAGATTTGGAATTTGTGACAAACACTTCC 862
Qy 779 TTGTTTAAGAGATAGAAATTTGTAAGTCTACTTTTGGGACTAGGTTTCAGGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
Qy 839 ACAGTATTCCTGGGCTAGAGTTTCTTTCACAAAAGCTTTGCTGTTTATAACTCTTATAGA 898
Db 923 TCAATACTCTCAGGCTCGTGTGATGCTTGTAGACTATAGCAATGATTTCTATAGTAGA 982
Qy 899 TGACACTTATGATGCTGATGTTACTTATGAAAGACTTTAAGATCTTTACTGAAGCTGTTGA 958
Db 983 TGACACATTCGATGCTTATGGCATTTGCAAGAACTTGAGATCTACACCGATGCCATACA 1042
Qy 959 AAGGTGTCATTTACATGCTTAGACACACTTCCAGATATACGAAACCGATATACAAAT 1018
Db 1043 GAGGTGGGATTTAGCCAAATGATCGGCTCCCTGATTAATGAATACTAGTTTCAAAAGC 1102
Qy 1019 ATTCAAGGATATACACAGAAATGGAAGAAATTTCTTGAAGAGGAGGAAGAAACAGATCT 1078
Db 1103 ACTTTTATGATCTCTCAATGATTAATGGAATGAGTGTGTCACAGGATGTTAGTCTGATGT 1162
Qy 1079 ATTTAACTCGGCAAGAAATTTGTGAAGAGTTTGTGTAAGAACTGATGTTGTAAGCAAA 1138
Db 1163 TGTTCCTACTCGCAAGAAAGAAATGAAAGAAATCGTGAGAAACTATTTTGTGGAAGCAAA 1222

Qy 1139 ATGGCAAAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTTAATCATTTAC 1198
Db 1223 ATGTTTCATTTGAAGGATATATGCCGCCAGTCTCTGAGTATCTTAGCAATGCAATAGCTAC 1282
Qy 1199 TGGCGGTGCTAACTGCTTACAACTTTGTTATCTTTGGCATGAGTATATTTACAAA 1258
Db 1283 CAGCACTTATTTACTTGTCTTACGACTCATCTTATTTGGGCATGAAAGTCTCTGTAACAAGCA 1342
Qy 1259 AGAGTCTGTCGAATGGGCTGCTCTGCACCTCTCTTTTATAGATATCTCAGGTATACCTGG 1318
Db 1343 AGATTTTG---AATGGTTGGCAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
Qy 1319 TCGACGCTTAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACTATGAGGTTGAGAAGGGTAGAGGTCAAGTTGCCAC 1459
Qy 1379 GAGCTTTGAAAGTTATATGAAGGAATATAATGTCTCATGAGGATGATGCCAACCTTGAT 1438
Db 1460 TGGAAATTTGAAATTTTACATGAGAGATTTATGGTGTATCCACAGAAAAGGCAATGGAATAAT 1519
Qy 1439 TTACAAGAAAGTAGAAGATGTTGGAAGATATAAACCGAGAGTACCTCACAACTAAAAA 1498
Db 1520 CCAAGAAATGCTGAGACAGCATGGAAGATGTAATGAAGAAATCCTTCGACCACTCC 1579
Qy 1499 CATTCCAAGGCGCTTATTTGATGGCTGTGATCTATTTTGCCAGTTTCTTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTTCTCACTCGCATTTCTCAATCTTCTGCAATATCGATGTTACTTA 1639
Qy 1559 TGCAGGAA---AGGATAACTTCACAGTATGGGAGACGAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCAAAATCAAGATGATACACTCATCCGAAAAAGTACTAAAAACCTCATATTATTGTC 1699
Qy 1616 TCTACTCGTTTATCTCTATGAGTATATGA 1643
Db 1700 GTTCTTGGTGGACTCTATTGAAATTTAA 1727

RESULT 10

US-09-900-797-31
; Sequence 31, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-900-797-31

Query Match 20.0%; Score 330.4; DB 10; Length 1944;
Best Local Similarity 52.0%; Pred No. 4.5e-78;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

14 AGAAGAAAAACCTATTGCGCCCAATGCAACTTTCTCAAGCAATTTGGGGAGACAGTT 73
Db
92 AGAGGAGAGATTGTTGCGCCCATAGCTGCTTCTCTCCAGTCTTTGGGGTGAATGTTT 151
Qy
74 TCTCATCTATCAAAAGAGTAGAGCAAGGGTGAACAGATAGATGATTTAAAAA 133
Db
152 TCATTTCATCTCCCTCGCAATCAATGCTGGAANAATATGCTCAAGAGATCGAACTTT 211
Qy
134 AGAAGTGGGCAACTACTATAAAGAGCTTTGGATATTCCTATGAACAATGCTTGT 193
Db
212 GAAGGAACAATCAAGAAATATATATCTGCTATCTCTCGAAGAAATATGCTGAGAAAT 271
Qy
194 GAAGCTGATTGATGAATTCACGCTTTGGAATACCGTATCATCTTTGAACGGGAGATGA 253
Db
272 GGATCTGATAGACATTTGTTGAGCGCTTTGGCATTCCTTATCATTTGAAAAACAATAGA 331
Qy
254 TCATGATTCGAATGATTTATTAAGAAATATGCTGGAANAATATGCTCAAGAGATCGAACTTT 298
Db
332 TGATATGTTGGATCAATTTTAAAGAGAGATCTTAACCTTTGAGGCTACGAGTACAATGA 391
Qy
299 TGGTGACCGCTCTCTTATGCTTCCGCTTATGCGAAGCAAGATATATGTTTACATG 358
Db
392 TTACAAACTTTATCCGTTCAATTTGAGCTATTGAGACAAATGTTTACATATCTCCCC 451
Qy
359 TGATGTTTCAATAACTATAAAGCAAAATGAGCGTTTCAAGCAATCGTTAGTAAATGA 418
Db
452 AAAAATTTTATTAGATTCGAAGATGCAAAAGGCAAAATTTAAAGAAATCTCTTTGTAACA 511
Qy
419 TGTTGAGGTTGCTTGAATGTTAGAGCAAACTTCTATGAGGTTACCTGGGAGATAT 478
Db
512 CATCAAGGCTTTTGAATATACGAAGCTTCGATGTAAGGACTCATGGAAGAGATAT 571
Qy
479 ATTAGAAGATCTCTTGTGTTTACAGATCTGCTTATGATTAATGACAAAGAGTCTTT 538
Db
572 TTTGAGAGGCACTTGTCTTCTCTCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCT 622
Qy
539 TTCTACAAACCCGCTCTTTTACGGAATACAAAGGCACTTAAAGCAACCCCTTTGGAA 598
Db
623 ACATTTGAAGTCACTCTGAGTAAAGCAAGTGAACATGCTTGAAGCAATCTCTCCATA 682
Qy
599 AAGTTGCCAAGATAGAGGGGCGCAGTACATCTCTTCTATCAACAAGATCTCTCA 658
Db
683 GAGCATTTCAAGAGTTGAGACAGCTACTCTATCTATCTACGAGAGGAGGAGACAGAA 742
Qy
659 TAACAGACTTTTAACTTGAATGCTTGAATGAGTTCATTTGCTTCAATTTGCTTCACTT 718
Db
743 GAATGATGTTGCTTCAATTTGCAAACTGGAATTTCACTTCACTTCACTTCACTTCACT 802
Qy
719 GGAAGAGCTCAGCCATGTTGCAAAATGTTGGAAGCTTTGATATCAAGAAGACGCAAC 778
Db
803 ACAAGAACTTAGTGAAGTATCAAGTGTGGAAGATTTGGAATTTGTGACAACTTCC 862
Qy
779 TTGTTAAGAGATAGAAATGTTGAATGCTATTTTGGGAGCTAGGTTCAAGCTATGAGCC 838
Db
863 ATATGCTAGGATAGAGAGTGAATGCTATTTTGGAGATGCTGCTGCTGCTGCTGCTGCT 922
Qy
839 ACAGTATTCGGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTATTAACCTTATAGA 898
Db
923 TCAATACTCTCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 982
Qy
899 TGACACTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
Db
983 TGACACATTCGATGCTTATGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
Qy
959 AAGTGGTCAATTAATGCTTATGAGCACTTCCAGAAATACATGAACCGATATACAAAT 1018
Db
1043 GAGTGGGATATTAGCCAAATTTGATCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1102
Qy
1019 ATTATGATATACATACAGAAATGGAATTTTTCGAAAGGAGGAGGAGGAGGAGGAGGAG 1078
Db
1103 ACTTTAGATCTCTCAATGATTAATGAATGAGTGTCTCAAGGATGGTAGATCTGATGT 1162
Qy
1079 ATTTAACTGCGGCAAGAAATTTGTAAGAGATTTGTTAGAAACCTGATGTTGAAGCAAA 1138

Db
1163 TGTTCATCTACGCAAGAAAGAAATGAAAGAAATCGTGAGAAATATTTTGTGGAGCAAA 1222
Qy
1139 ATGGGCAATGAGGAGACATACCAACCACTGAAGAGACATGATCCAGTTGTAATCATTAC 1198
Db
1223 ATGTTTCAATTTGAAGATATATGCGCCAGTCTCTGAGTATCTTAGCAATGATTTAGCTAC 1282
Qy
1199 TGGCGGTGCTAACTGCTTCAACCACTGTTATCTTTGGCAATGATGATATATTACAAA 1258
Db
1283 CAGCACTTATTTACTTGTCTACGACTACATCTTATTTGGGCATGAAGTCTGCTAACAGCA 1342
Qy
1259 AGAGTCTGCGAATGAGGCTGCTCTGCACTCTCTCTTTTAGATATCTCAGGTATCTTGG 1318
Db
1343 AGATTTTG---AATGTTGGCCAAAGAACCCCTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
Qy
1319 TCGACCCCTAAATGATCTCATGACCCCAAGCGGAGCAAGAAAGAAACATAGTTTCATC 1378
Db
1400 CCGAGTCTATAGATGACATAGCCACCTATGAGTTGAGAGGGTAGAGTCAAGTTGCCAC 1459
Qy
1379 GAGCCTTGAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAACTTGTAT 1438
Db
1460 TGGAAATGGAATTTACATGAGAGATTTATGTTATCCACAGAAAGGCGCATGGAATAAT 1519
Qy
1439 TTCAAGAGAGTAGAGATGTTGGAAGATATAATAATGTTCAATGAGGAGTATGCCCAAACTTGTAT 1498
Db
1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATGTAATCAAGAAATCTCTTCGACCACTCC 1579
Qy
1499 CATTCAGAGCCGCTTATTTGATGCTGATCTATTTGTCAGTCTTTCTTGAAGTTCAATA 1558
Db
1580 CGTCTCTACAGAAATTTCTCTGCAATCTCAATCTTCTGCTCGCATTTATCGATTTACTT 1639
Qy
1559 TGCAGGAA---AGGATAACTTCACAGTATGGGAGAGCAATACAAACATCTCATAAAGTC 1615
Db
1640 TAAGCACAATCAAGATGATACACTCATCCGAAAGAAAGTACTAAAAACCTCATATTATTGC 1699
Qy
1616 TCTACTCGTTTATCTTATGATGATATGA 1643
Db
1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 11

US-09-893-820-31
; Sequence 31, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-893-820-31

Query Match		20.0%;	Score 330.4;	DB 13;	Length 1944;
Best Local Similarity		52.0%;	Pred. No. 4.5e-78;		
Matches 857;		Conservative	0;	Mismatches 761;	Indels 30; Gaps 4;
QY	14	AGAGAAACCTATTTCGCCCAATGCCAACTTCCTCCCAAGCAATTTGGGGAGATCAGTT	73		
Db	92	AGAGAGGAGATTTGTCGCCCAATGCTGCTCTCCAAAGTCTTTGGGGTATCGTTT	151		
QY	74	TCTCATCTCAAAAGCAGTAGAGCAAGGGTGGACAGATAGTGAATTTAAAAA	133		
Db	152	TCATTTCCTCCCTCGCAATTCAGATTCGCGAAATATGCTCAAGAGATCGAACTTT	211		
QY	134	AGAGTGGCGCACTACTAAAGAACTTTGGATATTCCTATGAAACATGCCAAATTTGTT	193		
Db	212	GAGGAAACATCAAGAAATATATATCTGATCTCTCGAAGACATTCGCTGAGAAAT	271		
QY	194	GAGCTGATGATGAATTAACAGCCCTTGAATACCGTATCACTTTGAACGGGAGATGA	253		
Db	272	GGATCTGATGACATTTGTTGAGCGCTTGGCATTCGCTTATCATTTTGAAGAAACAAATAGA	331		
QY	254	TCATGCTGCAATGATTTATGAACATATGTT-----GATGACTGGAA	298		
Db	332	TGATATGTTGGATCAATTTTCAAGCAGATCTTAATCTTTGAGGCTCACGAGTACAAATGA	391		
QY	299	TGGTGACCGCTCTTCTTATGTTTCCGTCCTTATGCGAAAGCAAGATATTTATGTTACATG	358		
Db	392	TTTACAACTTTATCGTTCAATTTTCGACTTTTGACACACATGTTTACATATCTCCCC	451		
QY	359	TGATGTTTCAATCACTAATAAGCAAAATGAGAGCGTTCACCAATTCGTTGCTGAATGA	418		
Db	452	AAAACTTTTATAGATTCGAAGATCAAAAGCAAAATTTAAAGAAATCTCTTTGTAACGA	511		
QY	419	TCTTGAAGGTTTGTGAGTTGACGAACCACTTCTATGAGGTTACCTGGGAGATTA	478		
Db	512	CATCAAGGCTCTTTGAACTTATCAAGGCTCGCATGTAAAGCACTCATGGAGAAATAT	571		
QY	479	ATTAGAAGTGTCTTTGGTTTACAGGATCTCTCTTTAGCATTTATGACAAAAGATGCTTT	538		
Db	572	TTTGGAGAGGACATGCTTTCTCTACTGCTCATCTTGAATCT-----CGAGCTCC	622		
QY	539	TTCTACAAACCCGGCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAA	598		
Db	623	ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCCCTTGAGCAATCTCTCCATA	682		
QY	599	AGGTTGCCAAGATAGAGCGCGCAGTACATCTCTTCTATCAACAAAGATTTCTCA	658		
Db	683	GAGCATTCGAAGTTGAGACAGCTACTCTCTCTATCTAGCAAGAGGAGAAACAGAA	742		
QY	659	TAAACAGACTTTACTTTAAACTTCTTAAGTTAGAGTTCAATTTGCTTTCAGTCAATTCACAA	718		
Db	743	GATGATGTTGCTTCAATTTGCAAACTGGACTTCACTTACTTACAGATTTGCACAA	802		
QY	719	GGAAGCTCAGCCATGTTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGCAGCC	778		
Db	803	ACAAGAACTTAGTGAAGTATCAAGGTGGTGGAAAGATTTGGATTTTGTGACAACTCTCC	862		
QY	779	TTGTTTAAAGATAGAAATTTGTAAGTCTACTTTTGGGACTAGGTTTCAGGCTATGAGCC	838		
Db	863	ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGGTGATGCTGAACC	922		
QY	839	ACAGATTTCCGGGCTAGAGTTTCTTTCACAAAGCTTTGCTGTTTATTAATCTTTATAGA	898		
Db	923	TCATACTCTCAGCTCGTGTGATGCTGCTAGACTATAGCAATGATTTCTATAGTAGA	982		
QY	899	TGACACTTATGATGCTATGTTTCTTATGAAAGCACTTAAGATCTTTACTGAAAGCTGTGA	958		
Db	983	TGACATTCGATGCTTATGCGCATTTGTCAAGAACTTGATCTACACCGATGCGCATACA	1042		
QY	959	AGGTTGCTCAATFACATGCTTACACACACTTCAGAAATACATGAACCGATATACAAAT	1018		
Db	1043	GAGGTGGATATTAGCCAAATGATCGGCTCCCTGATTAATGAAATCAGTTACAAAGC	1102		
QY	1019	ATTCATGGATACATACACAGAAATGGAAGAAATTTCTTGCAAAGAGGAGGAAACAGATCT	1078		

Db	1103	ACITTTAGATCTCTACAAATGATTAATGAATGGAGTTGTCGAAGGATGGTAGATCTGATGT	1162		
QY	1079	ATTTAACCTGGGCAAGAAATTTGTGAAAGAGTTTGTAGAAACCTGATGGTTGAAGCAAA	1138		
Db	1163	TGTTCACTACGCGAAGAAAGAAATCGTGAGAAATCTATTTTGTGGAAGCAAA	1222		
QY	1139	ATGGGCAAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAAATCATTTAC	1198		
Db	1223	ATGTTTCAATTTGAAGGATATATGCGCCAGTCTCTGAGTATCTTAGCAATGCAATAGCTAC	1282		
QY	1199	TGGCGGTGCTAACTGCTTACAAACAACTTTGTTATCTTGGCATGAGTGTATATTCACAA	1258		
Db	1283	CAGCACTTATTTACTTGTCTTACGACTACATCTTATTTGGGCATGAAAGTCTGCTTAAACAAGCA	1342		
QY	1259	AGACTCTGCGAATGGGCTGCTCTGACACCTCTCTTTTATAGATCTCAGGTATACTTCTG	1318		
Db	1343	AGATTTTGTG---AATGTTGGCCAAAGACCCCTTAAATCTTGAAGGCTAATGTGACGTTATG	1399		
QY	1319	TGACGCGCTAAATGATCTCATGACCCCAAGCGCCGAGCAAGAAAGAAACATAGTTTCATC	1378		
Db	1400	CCGAGTCAATAGATGACATAGCCACTATGAGTTGAGAAAGGTTAGAGGTCAGATTGCCAC	1459		
QY	1379	GAGCTTTGAAAGTTATATGAAGGAATATAATGTCTCAATGAGGAGTATGCCAAACCTTGAT	1438		
Db	1460	TGGAATTTGAATGTTTACATGAGAGATTTATGTTGTATCCACAGAAAGGCCATGGAATAAT	1519		
QY	1439	TTCAAGGAAGTGAAGATGTGGAAGATATAAACCGAGAGTACCTCACAACTAAATAA	1498		
Db	1520	CCAAGAAATGCTGAGACAGCATGGAAGGATGTAATGAAGGAATCCTTCGACCAACTCC	1579		
QY	1499	CATTCCAAAGGCGCTTATTTGATGGCTGTGATCTATTTGTGCCAGTTTCTTTGAAGTTCAATA	1558		
Db	1580	CGTCTCTCAGAGATTTCTCACTCGCATCTCAATCTTCTGCTCGCATTCGATGTTACTTA	1639		
QY	1559	TGCAGGAA---AGGATAACTTTCACAGTATGGGAGAGCAATACAAACATCTCATAAAGTC	1615		
Db	1640	TAGCACAATCAAGATGGATACACTCATCCGGAAGAAAGTACTTAAACCTCATATTATTGC	1699		
QY	1616	TCTACTCGTTTATCTCTATGATATGA	1643		
Db	1700	GTTCTGTTGGACTCTATTGAAATTTAA	1727		

RESULT 12

US-09-887-586A-1
; Sequence 1, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)...(1668)
US-09-887-586A-1

Query Match 19.4%; Score 320.6; DB 9; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

15 GAAGAAAACCTATTGCGCCCATGTCACACCTTCTCCAGCATTTGGGAGATCAGTTT 74
119
52 GAAGAAGAGATTGTTGCGCCCGTCCGCACTTCTCCCTAGTTTGGGGTGATCAGTTT 111
117
75 CTATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
133
112 CTTTCATCTCCATTA---AAATCAGGTTGCAGAAAGATAGTCAAGAGATTGAAGA 168
166
135 GAAGTGGCGCAACTACTAAAGAAAGCTTTGGATATTCCTATGAACATGCGCAATTTGTTG 194
192
169 TTGAAGCAACAAACGAGGAAATATGCTGTTAGCAACTGGAATGAAATTTGGCTGATACATG 228
225
195 AAGCTGATGATGAATTAACAGCGCTTGAATACCGTATACATTTGAACGGGAGATTGAT 254
251
229 AATTGTATGACACTATTGAACGCGCTTGGCATATCTTACCACCTTTGAGAAAGAAATTGAT 288
284
255 CATGCATTGCAATGATTTATGAACATATGTTGATTAATGTAATGGAATGGTGAACGCTCTCC 314
310
289 GATATTTTGGATCAGATTTTACCAAACTCAAACTGCAAGGATTTGTGCACTTCTGCA 348
344
315 TTATGTTTCCGCTTATGCGAAAGCAAGGATATTTATGTTACATGATGATTTTCAATAAC 374
370
349 CTTCAATTTTCGATTGCTCAGGCAACATGTTTCAACATCTCTCCGAAATTTTCAGCAA 408
403
375 TATAAGACAAAATGAGCGCTTCAAGCAATGTTAGCTATGATGTTGAAGTTTGCTT 434
430
409 TTCCAGACGAAATGGAATTTCAAGGATCTCTTGTCTAGTATGTTCTTAGGATTTATG 468
463
435 GAGTTGTACGAAGCAACTTCTATGAGGCTACTGGGAGATATATATAGAGATGCTCTT 494
490
469 AACTTGTATGAAGCTTCAATGTAAGGACTCATGCTGAGTATCTTGAAGACGCACTT 528
523
495 GGTTTTACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCCGCT 554
549
529 GCTTTTCTC-----CACTATCCATCTGAACTCTGAGCTCCACATTTGAAATCTCA 579
574
555 CTTTTCACGAAATACAAACGGGCACTAAGCAACCCCTTTGGAAAGGTTGCAAGATA 614
609
580 CTTAGGAGCAAGTGACACATGCGCTTGGCAATTTGTTGCAACAGGGTGTCTCTAGATC 639
634
615 GAGCGGCGCAGTACATTCCTT---TCTATCAACACAGATTTCTATCAACAGACTTTA 671
666
640 GAGACCGGATTTCTCATCTCATCAATCTATGACAAAGGAACTATGCAAGATATGTTGTTA 699
694
672 CTTAAACCTTGCTAAGTTAGAGTTCAATTTGCTTCACTGATGTCACAAAGGAGGCTCAGC 731
726
700 CTTGCAATTTGCAATTTGATTTCACTTCTCAGATGTTGCAACAAACAGAACTTTGCT 759
754
732 CATGTGCAAAATGGTGGAAAGCTTTGATATCAAGAAAGACGCACTTTGTTTAAGAT 791
786
760 CAAGTATCAAGGTTGGTGGAAAGATTGGAATTTGTTAAACAACTTCCATATGCTAGAT 819
814
792 AGAATTTGGAATGCTACTTTTGGGACTAGTTCCAGGCTATGAGCCACAGTATTTCCCGG 851
846
820 CGAGTAGTGAATGCTACTTTTGGGCATTAGAGTTTATTTTGGCCCTCAATCTCTCAA 879
874
852 GCTAGAGTTTCTTCAAAAAGCTTGTCTGTTATAACTCTTATAGATGACACTTATGAT 911
906
880 GCTCGGCTAGCTGTTTGAAGCCATATCAATGATTTGATTTGATGACACCTTTGAT 939
934
912 GCGTATGCTATGTAAGAACTTAAGATCTTTACTGAAGCTGTTTGAAGGTTGCAATT 971
966
940 GCTTACGGTACAGTTAAGAACTTGAAGCATACACAGATGTCATCAAAAGATGGGATATC 999
994
972 ACATGCTTAGACACATTCAGATATACATGAACCGGATATCAATTTATTCATGATACA 1031
1026
1000 AACGAATTTGATCGGCTTCTCTGATTTACATGAAATCAGTTTCAAAAGCTATTCTAGATCTC 1059
1054

1032 TACACAGAAATGGAAGAAATTTCTTGGCAAGAGGAGAAAGACAGATCTATTTAATCGCGC 1091
1096
1060 TACAGGATTTATGAAAGGAATTTGTCTAGTGGCGAAGATCTCATATTTGTCTGCCATGCA 1119
1114
1092 AAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAATGGCAATGAG 1151
1147
1120 ATAGAAGAATGAAGAAGTAGTAAGAAATTTAATGTCGAGTCAACATGTTTATTGAA 1179
1174
1152 GGACACATACCAACCACTGAAGAGAGATGATCCAGTTGTAATCATTTACTTGGGGTCTAAC 1211
1206
1180 GATATACGCCACTGTTTCTGATACCTAAGCAATGCATAGCAACTACACATATTAC 1239
1234
1212 CTGCTTACAAACAACTTTTATCTTGGCATGAGTGATATTTCAAAAAGAGTCTTCGAA 1271
1265
1240 TACCTCGCAGCAACATCGTATTTTGGGCATGAAGTCTGC---CACGGAGCAAGATTTTGAG 1296
1291
1272 TGGGCTGTCTCTGCACCTCTCTTTTATAGATCTCAGGTATCTTGGTCGAGCCCTAAT 1331
1326
1297 TGGTTGTCAAAGAAATCCAAAAATTTCTTGAAGCTAGTGTAAATTTATGTCGAGTTATCGAT 1356
1351
1332 GATCTCATGACCCACAAAGCGCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTCAAAAGT 1391
1386
1357 GACACAGCCACGTACGAGGTTGAGAAAGCAGGGGACAAATTTGCACTGGAATTCAGTGC 1416
1411
1392 TATATGAAGAAATATATGTCATGAGGAGTATGCCCAACCTTGATTTTCAAGGAAGTA 1451
1446
1417 TGCATGAGAGATTATGATATATCAACAAAGAGGCAATGGCTAAATTTCAAAATATGGCT 1476
1471
1452 GAGATGTTGGAAGAGATATAAACCGAGAGTCTCAACACTAAACATTTCCAGGCGC 1511
1506
1477 GAGACAGCATGGAAGAGATATTAATGAAGGACTTCTTAGGCCACTCCCGTCTCTACAGAA 1536
1531
1512 TTATTCATGCTGTGATCTATTTTGTGCGAGTTTCTTGAAGTTCAATATGCAAGAAA---G 1568
1563
1537 TTTTAACTTCTTATTTCTCAATCTTGTCTGCTGTTGTTGAGGTTACATATACACATCTA 1596
1591
1569 GATAACTTCAACGATGAGGAGAGCAATACAACATCTCAATAAGTCTCTACTCGTTTAT 1628
1623
1597 GATGGATACACTCATCCGAGAGAAAGCTTAAACCTCACATTTATTAACCTACTTTGGGAC 1656
1651
1629 CCTATGAGTATATGA 1643
1670
1657 TCCATCAAAATTTGA 1671

RESULT 13
US-09-903-012-1
; Sequence 1, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (25)....(1668)
US-09-903-012-1

Query Match      19.4%; Score 320.6; DB 9; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAAACCTATTGCGCCCAATTCGCCAATTTCTCCCAAGCATTTGGGGAGATCAGTTT 74
DB 52 GAAGAGAGATTTGTCGCCCGCTGCCGACTTCTCCCTAGTTTGGGGGTGATCAGTTC 111
QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAAAA 134
DB 112 CTTTCATTCTCCATTAA--AAATCAGGTTGCAGAAAAGTAGTCTCAAGAGATTGAAGCA 168
QY 135 GAAGTGGCGCACTACTAAAGAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTG 194
DB 169 TTGAAGGAAACAAACGAGGAATATGCTGTAGCAACTGGAATGGAATTTGGCTGATACACTG 228
QY 195 AAGCTGATTGATGAATTAACAGCCTTGAATACGATATCACTTTGAACGGGAGATTGAT 254
DB 229 AATTTGATAGACACTATTGAACGCTTGGCAATCTCTACCACTTTGAGAAAGAAATGAT 288
QY 255 CATGCATTGCAATGATTTATGAACATATGGTGATTAACCTGGAATGGTACCGCTCTCC 314
DB 289 GATATTTTGGATCAGATTTTACAACCAAACTCAAACTGCAACGATTTGTGCACTTCTGCA 348
QY 315 TTATGGTTCCGCTTATGCGAAGCAAGGATATTAATGTTACATGTGATGATTTCAATAAC 374
DB 349 CTTCAATTTTCGATTGCTCAGGCAACATGGTTTCAACATCTCTCTGAAATTTTCAGCAA 408
QY 375 TATAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTTAATGATCTTGAAGTTTGGCTT 434
DB 409 TTCCAAGACGAAATGGCAAAATCAAGAAATCTTCTGCTAGTGTCTTAGAATTTTG 468
QY 435 GAGTTGTACGAAGCAACTTCTATAGGGTACCTGGGGAGATATATATGAAGATGCTCTT 494
DB 469 AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTAGAAGCGCACT 528
QY 495 GGTTTTACAGATCTGCTTAGCATTTATGCAAAAGATGCTTTTCTCAACACCGCT 554
DB 529 GCTTTCTC-----CACTATCCATCTTGAATCTGACGCTCCACATTTGAATCTCCA 579
QY 555 CTTTATTACGAAATACACGGGCACCTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAATA 614
DB 580 CTTAGGGAGCAAGTACACATGCCCTTGAGCAATGTTTGCAACAGGCTGTTCTTAGATC 639
QY 615 GAGCGCGCGCAGTACATTCCTT---TCTATCAACAACAAGATTTCTCATACAAGACTTTA 671
DB 640 GAGACCGGATTTCTCATCTCATCAATCTATGCAAGGAACAATCGAAGAATAATGTGTTA 699
QY 672 CTTAACTTGTCTAAGTTAGATTCATTTGCTTCACTCAATTTGCAACAGGAGAGCTCAGC 731
DB 700 CTTGATTTGCCAAATTTGGATTTCAACTTGTCTCCAGATGTTGCAACAAACAAGAACTTGT 759
QY 732 CATGTGTGCAATGTTGGAAAGCTTTTCGATATCAAGAAGAACCCACCTGTTTAAAGAGAT 791
DB 760 CAAGTATCAAGTGTGGGAAAGATTTGATTTGTAACAACACTTCCATATGCTAGAGAT 819
QY 792 AGAATTTGTAATGCTACTTTTGGGAGCTAGTGTTCAGGCTATGAGCCACAGTATTTCCCGG 851
DB 820 CAGTAGTTGTAATGCTACTTTTGGGCATTAGGAGTTTATTTTGGAGCTCAATCTCTCAA 879
QY 852 GCTAGAGTTTCTTCAAAAAGCTTGTGTTTATTAACCTTTATAGATGACATTTATGAT 911
DB 880 GCTCGCGTCATGCTCGTTAAGACCATATCAATGATTTTCGATTCGATGACACCTTTGAT 939
QY 912 GCGTATGGTACTTATGAAGAACTTAAAGATCTTTTACTTGAAGCTGTGAAAGGTGGTCAATT 971
DB 940 GCTTACGGTACAGTTAAAGAACTTTAGGCAATACACAGATGCCATACAAAGATGGGATTC 999
QY 972 ACATGCTTAGACACACTTCCAGAATAACATGAACCGGATATACAAATTTATTCATGGATACA 1031
```

```
DB 1000 AACGAAATTGATCGGCTTCTGATTATCATGAAATCAGTTTACAAAGCTATTCTAGATCTC 1059
QY 1032 TACACAGAAATGAAGAAATTTCTTGCAAAGGAGGAAGACAGATCTATTTAACTGCGGC 1091
DB 1060 TACAGGATTTATGAAGAAATTTCTTAGTCCCGAAGATCTCATATTTGCTGCGCATGCA 1119
QY 1092 AAAGAATTTCTGAAAGAGTTTGTAGAAAACCTGATGTTGAAGCAAAATGGGCAAAATGAG 1151
DB 1120 ATAGAAGAATGAAGAGTAGTAAGAAATTAATGTCGAGTCAACATGGTTTATTGAA 1179
QY 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTCTGGCGGTCTTAAC 1211
DB 1180 GGATATACGCCACCTGTTCTGAATACCTAAGCAATGCATGACCACTTACCACATATTAC 1239
QY 1212 CTGCTTACAACTGTTTATCTTGGCATGATGATATATTTCAAAAGAGATCTCTCGAA 1271
DB 1240 TACCTCGGCAACATCGTATTTGGGCATGAAGTCTGC---CAGGACCAAGATTTTGAG 1296
QY 1272 TGGGCTGTCTCTGCACCTCTCTTTTAGATATCTCAGGTATCTTGGTTCGACGCTTAAT 1331
DB 1297 TGGTTGTCAGAAATCCAAAATTTCTTGAAGCTAGTGTAAATATATGTCGATTTATCGAT 1356
QY 1332 GATCTCATGCCACCAAGGCGGAGCAAGAAAGAAACATGTTTATCGAGCCTTGAAAGT 1391
DB 1357 GACACAGCCACGTACGAGGTTGAGAAAAGCAGGGGACAAAATTGCAACTGGAATTTGAGTGC 1416
QY 1392 TATATGAGGAATATATGTCATGAGGAGTATGCCCAACCTTGATTTACAAGGAAGTA 1451
DB 1417 TGCATGAGATTTATGATATATCAACAAAGAGGCAATGGCTAAATTTCAAAATATGGCT 1476
QY 1452 GAAGATGTTGGAAAGATATAAACCAGAGAGTACCTCAACATAAAACATTTCCAAAGGCG 1511
DB 1477 GAGACAGATGGAAGATATTAATGAAGGACTTCTTAGGCCCATCTCCGCTCTTACAGAA 1536
QY 1512 TTATGATGCTGTGATCTATTTGTGCGAGTTTCTTGAAGTTCAATATGCGAGAAA---G 1568
DB 1537 TTTTAACTCTATCTCAATCTTGTCTCGTATTTGAGGTTTACATATATACAAATCTA 1596
QY 1569 GATAACTTTCACCTATGCGGAGAGCAATACAAACATCTCAAAAGTCTCTACTCGTTTAT 1628
DB 1597 GATGATACACTCATCCGAGAAAGTCTTAAACCTCACATTTTAACCTACTTCTGTGGAC 1656
QY 1629 CCTATGATATATGA 1643
DB 1657 TCCATCAAAATTTGA 1671

RESULT 14
US-09-900-797-1
; Sequence 1, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)....(1668)
US-09-900-797-1

Query Match      19.4%; Score 320.6; DB 10; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAAACCTATTTCGCCCCATTGCAACTTTCCTCCCAAGCAATTTGGGAGATCAGTTT 74
DB 52 GAAGAGAGATTGTCGCCCCCGTCGCCGACTTCTCCCTAGTTTGTGGGTGATCAGTTT 111
QY 75 CTCTCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
DB 112 CTCTCTATCTCCATTAA---AAATCAGGTTGCAGAAAAGTATGCTCAAGAGATTGAAGCA 168
QY 135 GAAGTGGCGCACTACTAAAAGAGCTTTGGATATTCCTATGAAACATGCCAATTTGTTG 194
DB 169 TTGAAGGAACAAAACGAGGAATATGCTTTAGCAACTGGAATGAAATTTGGCTGATACACTG 228
QY 195 AAGCTGATTGATGAATTCACCGCTTGAATACCGTATCATTGCAACGGGAGATTGAT 254
DB 229 AATTGATAGACACTATTGAAACGCTTTGGCAATCTCTACACTTTGAGAAGAAATTGAT 288
QY 255 CATGCAATTGCAATGATTTATGAACATATGTTGATACTGGAATGGTGACCCGCTCTTCC 314
DB 289 GATATTTTGGATCAGATTTTACAAACCAAACTCAAACTGCAACGATTTGTGCACCTTCTGCA 348
QY 315 TTATGGTTCGCTTATGCGAAGAGGATATTTATGTTACATGATGATTTTCAATAAAC 374
DB 349 CTTCATTTTCGATTGCTCAGGCAACATGTTTCAACATCTCTCCTGAAATTTTTCAGCAA 408
QY 375 TATAAGACAAAATGGAGCGTTTCAAGCAATGTTAGTAAATGATGTTGAAGGTTTGCCT 434
DB 409 TTCCAAGACAAAATGGCAAAATTCAGGAATCTCTTGTAGTATGATCTTTAGGATTTATG 468
QY 435 GAGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATATATATAGAAGATGCTCTT 494
DB 469 AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTAGAAGACGCAT 528
QY 495 GGTTTTACAGATCTGCTTATGATATGACAAAAGATGCTTTTCTCAAAACCCGCT 554
DB 529 GCTTTTCTC-----CACTATCCATCTTGAATCTGACGCTCCACATTTGAAATCTCCA 579
QY 555 CTTTTACCGAATACACGGGCACCTAAGCACCCCTTTGGAAGAGGTTGCCAAGATA 614
DB 580 CTTAGGAGCAAGTGACATGCCCCCTTGAGCAATGTTTGCACAAAGGTTTCTCTAGATC 639
QY 615 GAGCGCGCGCAGTACATTCCTT---TCTATCAACAAGATTTCTCATACAAGACTTTA 671
DB 640 GAGACCCGATTCTTCATCTCATCAATCTATGACAGGAACAATCGAAGATAATGTGTTA 699
QY 672 CTTAAACTTGTAGTAGTTCATTTGCTTTCAGTCTATGACAAAGGAGAGCTCAGC 731
DB 700 CTTCCGATTGCAAAATGGATTTCACCTTGTCTCAGATGTTGACAAAACAAGAACTTGT 759
QY 732 CATGTTGCAATGTGGGAAGCTTTCGATATCAAGAAGACCCACTTGTTTTAAAGAT 791
DB 760 CAAGTATCAAGTGTGGGAAGATTGGATTTTGTAAACAACACTTCCATATGCTAGAT 819
QY 792 AGAATTGTTGAATGCTATTTTGGGACTAGTGTTCAGGCTATGAGCCACAGTATTCGCG 851
DB 820 CGAGTAGTTGAATGCTACTTTTGGGCAATAGAGTTTATTTTGGAGCTCAATCTCTCA 879
QY 852 GCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTAACTCTTATAGATGACACTTATGAT 911
DB 880 GCTCCGCTCATGCTCGTTAAGACCATATCAATGATTTGATGTTGATGACACTTTGAT 939
QY 912 GCGTATGTTACTTATGAAGACTTATAGATCTTATGAGAGCTGTTGAAGGTTGCTCAATT 971
DB 940 GCTTACCGGTACAGTTAAAGAACTTGAGGCATACACAGATGCCATACAAAGATGGGATC 999
```

RESULT 15

```
US-09-893-820-1
; Sequence 1, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
```


940 GCTTACGGTACAGTTAAAGAACTTGGGCAATACAGATGCCATACAAAGATGGGATATC 999
972 ACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA 1031
1000 AACGAAATTTGATCGCTTCTCTGATTTACATGAATAATCAGTTTACAAAGCTATTTCTAGATCTC 1059
1032 TACACAGAAATGGGAAGAATTTCTTCCAAAGGGAAGAAACAGATCTATTTAACTCGGGC 1091
1060 TACAGGATTTATGAAGAAATTTGTCTAGTCCGGAAGATCTCATATTTCTGTGCCATGCA 1119
1092 AAAGAAATTTGGAAGAGTTTGTAGAAACCTGTGTTGAAGCAAAATGGGCAAAATGAG 1151
1120 ATAGAAAGATGAAGAAAGTAGTAAGAAATTAATATGCTGAGTCAACATGTTTATGAA 1179
1152 GGACACATACCAACCACTGAAGAGCATGATCAGTTGTAATCATTACTCGCGGTGCTAAC 1211
1180 GGATATACGCCACCTGTTTCTGAAATACCTAAGCAATGCACTAGCAACTACCAATATAC 1239
1212 CTGCTTACAAACAATTGTTTATCTTGGCATGATGATATATTCACAAAGAGTCTGTCGA 1271
1240 TACCTCGGACAAACATCGTATTTGGGCATGAAGTCTGC--CACGGAGCAAGATTTTGAG 1296
1272 TGGGCTGTCTGCACTCTCTTTTATAGATACTCAGGTATATCTTGGTGCAGCGCTAAAT 1331
1297 TGGTTGTCAAGAAATCCAAAATTTCTTGAAGCTAGTGTATATATGTCGAGTTATCGAT 1356
1332 GATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCATCGAGCCCTTGAAAGT 1391
1357 GACACAGCACGTCAGAGTTTGAGAAAGCAGGGGACAAATTCGAATCGAAATTTGAGTGC 1416
1392 TATATGAAGGAATATAATGTCAATGAGGATATGCCAAACCTTGATTTTACAGGAAGTA 1451
1417 TGCATGAGAGATTTATGGTATATCAACAAAGAGGCAATGGCTTAAATTTCAAAATATGGCT 1476
1452 GAAGATGTGCGAAAGATATAAACCAGAGTACCTCAACAATAAAACATTTCCAAGGCCG 1511
1477 GAGACAGCATGGAAGATATTAATGAAGACTTCTTAGGCCCACTCCCGCTCTCTACAGAA 1536
1512 TTATTTGATGGCTGTGATCTATTTTGGCCAGTTTCTTGAAGTTTCAATATGCGAGGAAA--G 1568
1537 TTTTAACTCTCTATTTCTCAATCTTCTGCTCGTATTTGTTGAGGTTACATATATACAAATCTA 1596
1569 GATTAACCTTCAACGATGCGGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTAT 1628
1597 GATGGATACACTCATCCGGAGAAAGTCTTAAACCTCAATTTAAACCTTCTTGGGAC 1656
1629 CCTATGATATATGA 1643
1657 TCCATCAAAATTTGA 1671

Search completed: June 7, 2004, 14:35:01
Job time : 762 secs

Db Qy
940 GCTTACGGTACAGTTAAAGAACTTGGGCAATACAGATGCCATACAAAGATGGGATATC 999
Qy
972 ACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA 1031
Db
1000 AACGAAATTTGATCGCTTCTCTGATTTACATGAATAATCAGTTTACAAAGCTATTTCTAGATCTC 1059
Qy
1032 TACACAGAAATGGGAAGAATTTCTTCCAAAGGGAAGAAACAGATCTATTTAACTCGGGC 1091
Db
1060 TACAGGATTTATGAAGAAATTTGTCTAGTCCGGAAGATCTCATATTTCTGTGCCATGCA 1119
Qy
1092 AAAGAAATTTGGAAGAGTTTGTAGAAACCTGTGTTGAAGCAAAATGGGCAAAATGAG 1151
Db
1120 ATAGAAAGATGAAGAAAGTAGTAAGAAATTAATATGCTGAGTCAACATGTTTATGAA 1179
Qy
1152 GGACACATACCAACCACTGAAGAGCATGATCAGTTGTAATCATTACTCGCGGTGCTAAC 1211
Db
1180 GGATATACGCCACCTGTTTCTGAAATACCTAAGCAATGCACTAGCAACTACCAATATAC 1239
Qy
1212 CTGCTTACAAACAATTGTTTATCTTGGCATGATGATATATTCACAAAGAGTCTGTCGA 1271
Db
1240 TACCTCGGACAAACATCGTATTTGGGCATGAAGTCTGC--CACGGAGCAAGATTTTGAG 1296
Qy
1272 TGGGCTGTCTGCACTCTCTTTTATAGATACTCAGGTATATCTTGGTGCAGCGCTAAAT 1331
Db
1297 TGGTTGTCAAGAAATCCAAAATTTCTTGAAGCTAGTGTATATATGTCGAGTTATCGAT 1356
Qy
1332 GATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCATCGAGCCCTTGAAAGT 1391
Db
1357 GACACAGCACGTCAGAGTTTGAGAAAGCAGGGGACAAATTCGAATCGAAATTTGAGTGC 1416
Qy
1392 TATATGAAGGAATATAATGTCAATGAGGATATGCCAAACCTTGATTTTACAGGAAGTA 1451
Db
1417 TGCATGAGAGATTTATGGTATATCAACAAAGAGGCAATGGCTTAAATTTCAAAATATGGCT 1476
Qy
1452 GAAGATGTGCGAAAGATATAAACCAGAGTACCTCAACAATAAAACATTTCCAAGGCCG 1511
Db
1477 GAGACAGCATGGAAGATATTAATGAAGACTTCTTAGGCCCACTCCCGCTCTCTACAGAA 1536
Qy
1512 TTATTTGATGGCTGTGATCTATTTTGGCCAGTTTCTTGAAGTTTCAATATGCGAGGAAA--G 1568
Db
1537 TTTTAACTCTCTATTTCTCAATCTTCTGCTCGTATTTGTTGAGGTTACATATATACAAATCTA 1596
Qy
1569 GATTAACCTTCAACGATGCGGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTAT 1628
Db
1597 GATGGATACACTCATCCGGAGAAAGTCTTAAACCTCAATTTAAACCTTCTTGGGAC 1656
Qy
1629 CCTATGATATATGA 1643
Db
1657 TCCATCAAAATTTGA 1671

Search completed: June 7, 2004, 14:35:01
Job time : 762 secs

Query Match 19.4%; Score 320.6; DB 13; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;
Qy 15 GAAGAAACCTATTTCGCCCATTTGCCAACTTCTCCCAAGCAATTTGGGAGATCAGTTT 74
Db 52 GAAGAAGAGATTTGTCGCCCGCTGCGCGACTTCTCCCTAGTTTGTGGGGTATCAGTTT 111
Qy 75 CTCATCTATCAAAAGCAAGTAGACAGAGGCGTGGCAACAGATAGTGAATGATTTAAAAAAA 134
Db 112 CTTTCAATCTCCATTAA--AAATCAGTTTGCAGAAAGTATGCTCAGAGATTTGAACA 168
Qy 135 GAAGTGGCGCAACTACTAAAGAAAGTTTGGATATTCCTATGAACAAATGCCAATTTGTTG 194
Db 169 TTGAAGAAACAAACGAGGAATATGCTTTAGCAACTGGAATGAATTTGGCTGATACACTG 228
Qy 195 AAGCTGATTTGATGAATTTCAACGCTTGGATACCGTATACCTTGAACGGGAGATTTGAT 254
Db 229 AATTTGATAGACACTATTGAACGCTTGGCATATCTTACCACCTTGAAGAAAGAAATTTGAT 288
Qy 255 CATGCAATTCGAAATTTATGAAACATATGTTGATAAATCTGGAATGTTGACCGCTCTTCC 314
Db 289 GATATTTTGGATCAGATTTTACACCAAACTCAACTGCAACGATTTTGTGCACTTCTGCA 348
Qy 315 TTATGTTTCCGCTTTATGCGAAAGCAAGGATATATATGTTTACATGATGATTTTCAATPAC 374
Db 349 CTTCAATTTTCGATTTGCTCAGGCAACATGTTTCAACATCTCTCCGAAATTTTCAGCAA 408
Qy 375 TATTAAGACAAATAGGAGGTTTCAAGCATCTTACGATCTTACGATTTGAGGTTTCTT 434
Db 409 TTCCAAGAGCAAAATGGCAAAATTCAGGAATCTCTTCTAGTATGTTCTTAGGATTTATG 468
Qy 435 GAGTTGTAGCAACCACTTCTATGAGGATACCTGGGAGATATATATGAGAGTGTCTCTT 494
Db 469 AACTTGTATGAGCTTCACTATGTAAGGACTATGCTGACGATATCTTAGAAGCGCACTT 528
Qy 495 GGTTTTACAGATCTCGCTTTAGCATATGACAAAGATGCTTTTCTCAAAACCCGCT 554
Db 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCAGTCCACATTTGAAATCTCCA 579
Qy 555 CTTTATCCGAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAGGTTGCCAAGATA 614
Db 580 CTTAGGGAGCAAGTGAACATGCCCTTGAGCAATGTTTGCAAGGGGTGTTCTTAGAGTC 639
Qy 615 GAGGCGGCGCAGTACATTCCTT---TCTATCAACAAAGATTTCTATAACAAGACTTTA 671
Db 640 GAGACCGGATCTTCACTCATCAATCTATGACAAAGCAATCGAAGATATATGCTTGA 699
Qy 672 CTTAAACTTCTAAGTTAGATTTCAATTTCTCAGTCATTTGCAACAAAGGAGAGCTCAGC 731
Db 700 CTTTCGATTTGCCAAATTTGGATTTCACTTCTCCAGATTTGCAACAAAGAACTTGTCT 759
Qy 732 CATGCTGCAATGTTGGAAGCTTTTCGATATCAAGAGAACGCCACCTTTTAAAGAT 791
Db 760 CAAGTATCAAGGTTGGTGAAGATTTGATTTTGTAAACACACTTCATATGCTAGAGAT 819
Qy 792 AGAATTTGTAATGCTACTTTTGGGAGTACTAGTTTCAAGGCTATGAGCCACAGTATTTCCCGG 851
Db 820 CGAGTAGTTGAATGCTACTTTTGGGCATTTAGGAGTTTATTTGAGCCCTCAATCTCTCA 879
Qy 852 GCTTAGAGTTTCTTCAAAAGCTTTGCTGTTATTAACCTTTATAGATGACACTTATGAT 911
Db 880 GCTCGGCTCATGCTCGTTAAGACCATATCAATGATTTTCGATTTGCGATGACACCTTTGAT 939
Qy 912 GCGTATGGTACTTTATGAGAACTTTAGATCTTTACTGAAGCTGTTGAAAGGTGCTCAAT 971

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 08:55:14 ; Search time 4504 Seconds
(without alignments)
10933.112 Million cell updates/sec

Title: US-09-763-822a-13

Perfect score: 1649

Sequence: 1 ccatgcactacagaaga.....ctatgatatatgagatcc 1649

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estio.*

8: em_estc.*

9: gb_estl.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394.2	23.3	749	13	BQ864711 QGC27H08.
2	351	21.3	692	13	BQ861631 QGC19C16.
3	341.8	20.7	699	13	BQ865374 QGC4a04.Y
4	325.2	19.7	697	13	BU027611 QHG6K24.Y

5	269.2	16.3	814	14	CF211234
6	256.6	15.6	742	14	CB970399
7	246.6	15.0	634	14	CB913201
8	244	14.8	779	14	CD486420
9	243.2	14.7	1981	11	AY105371
10	239.8	14.5	628	9	AU293055
11	239.2	14.5	762	12	BG645114
12	235.6	14.3	753	14	CB968730
13	232.6	14.1	663	10	AW685590
14	230.6	14.0	777	14	CD485858
15	226.4	13.7	807	12	BG588994
16	223.8	13.6	778	12	BM814641
17	221.4	13.4	638	10	AW559431
18	218	13.2	798	12	BG589051
19	214	13.0	627	9	AI727323
20	211.6	12.8	646	9	AI727526
21	211.2	12.8	496	14	CF609224
22	205	12.4	594	14	CB891633
23	204.8	12.4	608	9	AI727235
24	203.6	12.3	665	10	BF649458
25	194.4	11.8	638	10	AW584154
26	194	11.8	719	10	BE205416
27	191.6	11.6	778	12	BI421872
28	189.8	11.5	544	9	AU293056
29	189.4	11.5	763	14	CB969330
30	187.6	11.4	620	13	BU039172
31	185.6	11.3	624	9	AI730880
32	185.2	11.2	978	14	CK287078
33	181.2	11.0	593	10	BE449171
34	180.4	10.9	775	12	BG645053
35	178.8	10.8	534	14	CF605146
36	178.8	10.8	701	12	BI436194
37	178	10.8	611	12	BM406386
38	175.4	10.6	674	14	CA929582
39	174.2	10.6	785	12	BG645061
40	173.2	10.5	616	10	BF004127
41	172.8	10.5	2028	11	AY110842
42	171.6	10.4	536	10	AW616920
43	168.8	10.2	502	13	BQ104308
44	168.6	10.2	902	14	CK288631
45	168.4	10.2	628	10	BE919617

ALIGNMENTS

BQ864711 749 bp mRNA linear EST 14-AUG-2002
QGC27H08.yg.abi_QG_ABCDI lettuce salinas Lactuca sativa cDNA clone
QGC27H08, mRNA sequence.

ACCESSION BQ864711
VERSION BQ864711.1
KEYWORDS GI:22250176
SOURCE EST.

ORGANISM Lactuca sativa

Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.

REFERENCE 1 (bases 1 to 749)

AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.

Letture and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
Unpublished (2002)

JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig2363, see http://cgdb.ucdavis.edu/
for details.
Plate: QGC27 row: H column: 08.

FEATURES
source
1. .749
Location/Qualifiers
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGC27H08"
/lab_host="E.coli"
/clone_lib="QG ABCDI lettuce salinas"
/note="Vector: pBRCDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE="chemical induction"
TAG_LIB="QG ABCDI lettuce salinas"
TAG_SEQ="TGAGCCGGG"

ORIGIN
Query Match 23.3%; Score 384.2; DB 13; Length 749;
Best Local Similarity 70.3%; Pred. No. 4.2e-86;
Matches 515; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 109 AACAGATAGTGAATGATTAAAGGAGGCGGCACTACTAAAGAGGCTTTGGATA 168
DB 17 ATCGAATAGTCCATGATCTGAAGGAAGGTGAGAAAGATATCTGCTGCTTTATG 76
QY 169 TTCTATGAAACATGCCAATTTGTTGAAGCTGATGTAATCAAGCCCTTGAATAC 228
DB 77 TTCCAAATGGAACATACAAAACCTCTTGAACAAATCGACACAAATCCAGCTCTTGGATAG 136
QY 229 CGTATCACTTTGACGGGAGATGATCATGATGCAATGATTTATTAAGACATATGTTG 288
DB 137 CCTACTATTTCCGACCAAGAGATTAAGCAGGCTTCGAGCATATCTATGTGAATATGTTG 196
QY 289 ATAACTGAAATGGTGACCGCTCTTCCTTATGGTCCGCTTTATGCGAAGCAAGGATATT 348
DB 197 ATAACTGAGTGGTGGTAGCTCTTCTGTTTGGCTCATGCGACACAAAGGTTCT 256
QY 349 ATGTTACATGATGTTTTCATAAATACTATAAGACAAAATAATGAGCGTTCAAGCAATCGT 408
DB 257 ACGTTTCATGTGATATTTTCAATAACTACAAAAAAGAAATGGAGCTTTTAAGGAATCGT 316
QY 409 TAGCTAATGATGTTGAAGTGTGCTGAGTGTGACGAGCACTCTATGAGGTTACCTG 468
DB 317 TGACCAATGATCATCCATGAGATGCTTGAGTTATATGAGGCAACATTTATGAGAGTGAAG 376
QY 469 GGGAGATATATTAGAAGATGCTCTTGTTTTPACACGATCTCGTCTTACATTTACACAA 528
DB 377 GAGAGTTGTACTAGACGAGCTCTCTTTTTCACAAAACCTCATCTTGAAGAACTAGCAA 436
QY 529 AAGATGCTTTTTCACAAACCCCGCTCTTTTTCACGAAATAACACGGGCACCTAAAGCAAC 588
DB 437 AGGATCCCTGTTGGGTAACTCTACTCTTCAATACACATACAGGAGACACTAAAAACATC 496
QY 589 CCCTTTGAAGAAGTGTGCAAGATAGAGGCGCGCAGTACATTCCTTTCTATCAACAC 648
DB 497 CCATATTGAGAAGGTTGCAAGCTGAGGCAATACATACATTCCTTTCTACCAAAAAC 556
QY 649 AAGATTCTCATAAACAGCTTTTACTTAAACTGCTTAAGTTAGAGTTCAATTTGCTTCAGT 708
DB 557 AAGCTTCTTGTAAGTCTTTTACTTCACTTTTCAAGTTAGGTTCAATCTACTTCAAT 616
QY 709 CATTGCACAAAGGAGGTTCAGCCATGTCGCAATGTGTGAAATGTTGGAAAGCTTTTCGATACAGA 768

DB 617 CCTACACAAAGAAAGAGCTTAGTGAACCTTCCAGTGGTGAAGGTTTGTATGTACCGA 676
QY 769 AGAAGCGACCTTGTGTTAAGAGATAGAAATGTTGAATGCTACTTTTGGGAGCTAGGTTTCA 828
DB 677 ACAAGATACCTTATACACGGGACCGATTGTTGAATGTTACTTTTGGGCATTAGTGTAT 736
QY 829 GCTATGAGCCACA 841
DB 737 ACTTGTAGCCACA 749

RESULT 2
BQ861631
LOCUS
DEFINITION
QGC19C16.YG.ab1 QG ABCDI lettuce salinas Lactuca sativa cDNA clone
QGC19C16, mRNA sequence.
ACCESSION
BQ861631
VERSION
BQ861631.1 GI:22247096
KEYWORDS
EST.
SOURCE
Lactuca sativa
ORGANISM
Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.
1 (bases 1 to 692)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,X., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://comgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig2363, see http://cgdb.ucdavis.edu/
for details.
Plate: QGC19 row: C column: 16.
Location/Qualifiers
1. .692
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGC19C16"
/lab_host="E.coli"
/clone_lib="QG ABCDI lettuce salinas"
/note="Vector: pBRCDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE="chemical induction"
TAG_LIB="QG ABCDI lettuce salinas"
TAG_SEQ="TGAGCCGGG"

ORIGIN
Query Match 21.3%; Score 351; DB 13; Length 692;
Best Local Similarity 69.8%; Pred. No. 1e-77;
Matches 474; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 109 AACAGATAGTGAATGATTAAAGGAGGCGGCACTACTAAAGAGCTTTGGATA 168

Db 14 ATCGAATAGTCCATGATCTGAAAGGAAAGTGAGAAAGATATATCTGGCTGCTTTATATG 73

QY 169 TTCTTATGAACATGCCAATTTGTTGAGCTGATGATGAATTCACAGCCTTGAATAC 228

Db 74 TTCCATGGGAACATACAAAACCTCTTGAACATTAATCGACACATCAACGCTCTTGAATAG 133

QY 229 CGTATCACTTGAACGGGAGATGATCATGATGCAATGCAATGATTTATGAAACATATGCTG 288

Db 134 CCTACTATTTTCGACCAAGAGATTAAGCAAGCTTGCAGCATATCTATGTGAATATGCTG 193

QY 289 ATAACTGGAATGAGTACCGCTCTTCTTATGATGCTGCTTATGCGAAGCAAGATAT 348

Db 194 ATAACTGGAGTGGTGTAGCTCTTCTGTTTGGTTCGACTCATGCGACACAAAGGTTCT 253

QY 349 ATGTTACATGATGTTTCAATTAACATATAAGACAAAATGGAGCTTCAAGCAATCGT 408

Db 254 ACGTTTCACTGATATTTTCAATACTACAAAAGAAATGGAGCTTTTGAAGATCGT 493

QY 409 TAGCTAATGATGTTGAAGGTTTCTGTTGAGTGTGAGTATATGAGGCAACATTTATGAGTGAAG 373

Db 469 GGGAGATTAATGAGAGATGCTCTTGGTTTACAGATCTGCTTCTAGCATATGACAA 528

Db 374 GAGAAGTTTCTAGACGAGCTCTTCTTTTCAAAAACCTCATCTTGAATACTAGCAA 433

QY 529 AGATGCTTTTCTACAAACCCGCTCTTTTACGAAATACAAAGGCTTCAAGCAATC 588

Db 434 AGGATCTGTTTCGGTGTAACTTCTCTTCAATACACATACAGGACACCTAATAACATC 493

QY 589 CCCTTTGAAAGGTTGCCAAGAAATAGAGCGCGAGTACATTCCTTTCTATCAACAC 648

Db 494 CATATGAGAGGTTGCCAAGACTTGGAGCAATTACACTCATCTTCTTACCAAAAAC 553

QY 649 AGATTTCTCATACAGACTTTTACTTAAGTTCGTAAGTGTAGAGTTCAATTTGCTTCA 708

Db 554 AAGCTTCTTGTATGAGTCTTTTACTTCACTTTCAAAAGTTAGGTTCAATCTTCAAT 613

QY 709 CATGCAAGAGAGAGCTAGCAGATGTCGCAATGCTGGAAGCTTTCGATATCAAGA 768

Db 614 CCCTACACAAAGAGCTTAGTGAACCTTTCCAAAGTGTGGAAGGTTTGTATGATACCGA 673

QY 769 AGAAGCGACCTTGTTTAAG 787

Db 674 AACGATACCTTATACAG 692

RESULT 3

BO865374

LOCUS

DEFINITION

QGC4a04.Y9.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone

QGC4a04, mRNA sequence.

BO865374

BO865374.1 GI:22250839

EST.

KEYWORDS

SOURCE

ORGANISM

Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.

1 (bases 1 to 699)

Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Siabag, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compnoms.ucdavis.edu/

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akosika@ucdavis.edu [michelmore@vegmail.ucdavis.edu]

belongs to contig QG_CA_Contig2363, see http://cgdb.ucdavis.edu/ for details.

Plate: QGC4 row: a column: 04.

FEATURES source

Location/Qualifiers

1..699

/organism="Lactuca sativa"

/mol_type="mRNA"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="QGC4a04"

/lab_host="E.coli"

/clone_lib="QG_ABCDI lettuce salinas"

/note="Vector: pBRCDNA51AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/TAG_TISSUE=chemical induction

TAG_lib=QG ABCDI lettuce salinas

TAG_SEQ=GTAGCCGGG

ORIGIN

Query Match 20.7%; Score 341.8; DB 13; Length 699;

Best Local Similarity 69.5%; Pred. No. 2.2e-75;

Matches 477; Conservative 0; Mismatches 208; Indels 1; Gaps 1;

QY 109 AACAGATGATGATGATTTAAAAAGAGAGTGGGCAACTACTAAAGAGCTTGGATA 168

Db 14 ATCGAATAGTCCATGATCTGAAAGGAAAGTGAGAAAGATATATCTGGCTGCTTTATATG 73

QY 169 TTCTTATGAACATGCCAATTTGTTGAGCTGATGATGAATTCACAGCCTTGAATAC 228

Db 74 TTCCATGGGAACATACAAAACCTCTTGAACATTAATCGACACATCAACGCTTGAATAG 133

QY 229 CGTATCACTTGAACGGGAGATGATCATGCTTGAATGCTATTTATGAAACATATGCTG 288

Db 134 CCTACTATTTTCGACCAAGATATTAAGCAAGCTTGCAGCATATCTATGTGAATATGCTG 193

QY 289 ATAACTGGAATGAGTACCGCTCTTCTTATGATGCTTCTTATGCGAAGCAAGATAT 348

Db 194 ATAACTGGAGTGGTGTAGCTCTTCTGTTTGGTTCGACTCATGCGACACAAAGGTTCT 253

QY 349 ATGTTACATGATGTTTCAATACTATAAGACAAAATGGAGCTTCAAGCAATCGT 408

Db 254 ACGTTTCACTGATATTTTCAATACTACAAAAGAAATGGAGCTTTTGAAGATCGT 313

QY 409 TAGCTAATGATGTTGAAGGTTTCTGTTGAGTGTGAGTATATGAGGCAACATTTATGAGTGAAG 468

Db 314 TGACCAATGACATCCATGAGATGCTTGGTATATGAGGCAACATTTATGAGTGAAG 373

QY 469 GGGAGATTAATGAGAGATGCTCTTGGTTTACAGATCTGCTTCTAGCATATGACAA 528

Db 374 GAGAAGTTTCTAGACGAGCTCTTCTTTTCAAAAACCTCATCTTGAATACTAGCAA 433

QY 529 AGATGCTTTTCTACAAACCCGCTCTTTTACGAAATACAAAGGCTTCAAGCAATC 588

Db 434 AGGATCTGTTTCGGTGTAACTTCTCTTCAATACACATACAGGACACCTAATAACATC 493

QY 589 CCCTTTGAAAGGTTGCCAAGAAATAGAGCGCGAGTACATTCCTTTCTATCAACAC 648

Db 494 CATATGAGAGGTTGCCAAGACTTGGAGCAATTACACTCATCTTCTTACCAAAAAC 553

QY 649 AGATTTCTCATACAGACTTTTACTTAAGTTCGTAAGTGTAGAGTTCAATTTGCTTCA 708

Db 554 AAGCTTCTTGTATGAGTCTTTTACTTCACTTTCAAAAGTTAGGTTCAATCTTCAAT 612

DB	Sequence
Ddb	61 TAAACAAGGACCCCTTAGATGGAACTGTACTCTTTCTTCGATGCACATAGAGGAAG 120
Qy	577 CACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAATAGAGCGGCGCAGTACATCCCTT 636
Ddb	121 CACTAGAGCGACCCATATGGAAAAGATTGCCAAGGCTAGATACGGTGCATACATACCTT 180
Qy	637 TCTATCAACAACAGATTCTCATACAAGACTTTACTTAAACTTCTTAAGTTAGAGTTCA 696
Ddb	181 TCTATGAACACGACAGATTTCATATGAATCGCTAGTAGACTCCGCGAAGTTTGGATTTC 240
Qy	697 ATTTCCTTCAGTCATTGTCACAGGAAGAGCTCAGCCATGCTGTGCAAAATGGTGGAAAAGCTT 756
Ddb	241 ACAGGCTTCAGTCATTGTCACAAAAGAGGCTTAGCCCACTTGCCTCAAGTGGTGGAAAAGCTT 300
Qy	757 TCGATATCAAGAAGAACGCACCTGTTTAAAGAGATAGAAATGTTTGAATGCTACTTTTGGG 816
Ddb	301 TTGACCCCTCCAAAGAATCTGCAATTATGTAAGAGAGACAGAAATGTTGAAGCATATTTCTGGA 360
Qy	817 GACTAGGTTTCAGGCTATGAGCCACAGTATTCGGGGCTAGAGTTTCTTCACAAAAGCTG 876
Ddb	361 TAGTGGGTGTTCTACTTTGAGCCTCAATATTTCTGGTCCAGAAATTTCTTAACAAAAAATGT 420
Qy	877 TTGCTGTTTAACTCTTATAGATGACACTTATGATGCGTATGCTACTTATGAAGAAGCTTA 936
Ddb	421 TTAATAATGTTATCAATTTTAGATGACACTTATGATGCTTATGCTTTATGAGAAGCTTG 480
Qy	937 AGATCTTTACTGAAGCTGTGAAAGGTGGTCAATTACATGCTTAGACACACACTTCCAGAAT 996
Ddb	481 AGATTTTCACCAAGCAGTTCAAAGCTGGTCAATGCGCTGCATCGATGCACTTCTCTGATT 540
Qy	997 ACATGAACCGATATACAAATTTATTTCATGGATACATACACAGAAATGGAGAATTTCTTG 1056
Ddb	541 ACATGAAGCTGTGCATACAAAATCTCTTGGATTGTTTACCATGAATAGAGGAATAATGG 600
Qy	1057 CAAAGGAGGGAAGACAGATCTATTAACTGCGGCAAGAAATTTGTGAAAGAGCTTTGTTA 1116
Ddb	601 CAAACGAAGAAAAGCATATCAAGTTCCATGCCGAGAGTTGNATCAAGAGATGAGTA 660
Qy	1117 GAAACCTGATGTTGAAGCAAAATGGCAAAATGAGGG 1153
Ddb	661 GATGCTACATGATTGAAGCAAAATGGAGAACGACGG 697

RESULT 5

CF2111234

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

CF2111234 814 bp mRNA linear EST 01-AUG-2003

CAB20007_Iva_Fa_G05 Cabernet Sauvignon Flower bloom - CAB2 Vitis

vinifera cDNA clone CAB20007_Iva_Fa_G05 5', mRNA sequence.

CF2111234

CF2111234.1 GI:33405607

EST.

Vitis vinifera

Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.

1 (bases 1 to 814)

Goes da Silva,F., Iandolo,A., Lam,H., Baek,J., Jones,K. and Cook,D.

Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages

Unpublished (2003)

Contact: Douglas Cook, PhD

CAES Genome Facility

UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA

Tel: 530 754 6561

Fax: 530 754 6617

Email: drcook@ucdavis.edu

Seq primer: ACGGTACCGACATATGCC.

Location/Qualifiers

1..814

/organism="Vitis vinifera"

FEATURES

source

Db 680 CGGATGAACGAGTAGATTCTAGGAGGCTTTGATCAGCGATCTAAGAGCGCATGCTAGGCT 621
Qy 439 TGTACGAGCAACTTCTATGAGGACCTGGGAGATTATATAGAGATGCTCTTGGTT 498
Db 620 TGTATGAAGCTGCACATCTGAGGTTCTAGGAGGACATATCTGCAAGACACTTGTCT 561
Qy 499 TTACACGATCTCGTCTTACGATTAATGACAAAGAGTCTTTTCTACAAACCCCGCTCTTT 558
Db 560 TCACACCACTCATCTCAAGGCCATGTTAGAAAGTTAGGATAT-----CATCTTG 510
Qy 559 TTACGGAAATACACGGGACCTAAAGCAACCCCTTTGGAAAGGTTGCCAAGATAGAGG 618
Db 509 CAGAAACAGTTGCTCATGCCCTGAACCGGCCCATTAGAAAGGTTGGAGAGCTAGAG 450
Qy 619 CGGCGCAGTACATCTCTTCTATCAACACAGATTTCTATACAGACTTTTACTTAAAC 678
Db 449 CAAGATGTTATATCTGTCTACCAAGATGAAGCTTTCCATGATATAAACTTTTACTAGAGC 390
Qy 679 TTGCTAAGTTAGAGTTCAATTTGCTTCACTCATTTGCAAGGAAGAGCTTCAGCCATGTGT 738
Db 389 TAGCAAAATTAGATTTCAATCTAGTGCAGTCACTGCACAAAGGAGCTAAGCAATCTTG 330
Qy 739 GCAATGTTGAAAGCTTTTCGATATCAAGAAAGACGACCTTTGTTTAAAGATAGAAATG 798
Db 329 CAAGGTGTTGAAAGAAATTAGACTTTTGCTACAAAGTTTACCTTTTGACAGAGACAGATTGG 270
Qy 799 TTGAATGCTACTTTTGGGAGCTAGGTTTCAAGGCTATGAGCCACAGTATTCGCGGCTAGAG 858
Db 269 TTGAAGGCTACTTTTGGATGCTATGAGGCTATTTGAGCCCAATCTTACGCGGCTAGAC 210
Qy 859 TTTTCTTCAAAAAGCTTTGCTGTATTAACCTTTATAGATGACACTTTATGATGCGTATG 918
Db 209 GAATTTCAACCAAGTAATTTGCCATGACATCCATTTTATGATATCCATGATGCATATG 150
Qy 919 GTACTTATGAAGAACTTAAAGTCTTTTCTGAAAGCTGTTGAAAGGTTGCTCAATTACATGCT 978
Db 149 GCACCCCTGAAGAACTTCAAGCTCTTTCATAGAAGCCATTTGAGAGATGGGATATTAAACAGCA 90
Qy 979 TAGACACACTTCCAGAAATACATGAACCGATATACAAATTTTATGGAATACATACACAG 1038
Db 89 TAAATCAGCTTCCAGAAATACATGAAGAACTCTGCTATGTCGCACTTTAGATGTGTACAAAG 30
Qy 1039 AAATGGAAGAA 1049
Db 29 AAAAAAAAAA 19

RESULT 7
CB913201
LOCUS
DEFINITION
CB913201 634 bp mRNA linear EST 25-APR-2003
stressed berries of Vitis vinifera var. Chardonnay
CDNA clone VDL163B02 5, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CB913201.1 GI:30127862
EST.
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 634)
Cushman,J.C.

REFERENCE
AUTHORS
TITLE
An expressed sequence tag database for abiotic stressed berries of
Vitis vinifera var. Chardonnay
JOURNAL
COMMENT
Unpublished (2002)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650

Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 163 row: B column: 02
Seq primer: T3 20mer
High quality sequence stop: 634.
Location/Qualifiers
1. .634
/organism="Vitis vinifera"
/mol_type="mRNA"
/db_xref="taxon:29760"
/clone="VVD163B02"
/tissue_type="berries"
/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
/clone_lib="An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay"
/note="Vector: lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

FEATURES
source
Query Match 15.0%; Score 246.6; DB 14; Length 634;
Best Local Similarity 63.0%; Pred. No. 2.4e-51;
Matches 381; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
Qy 588 CCCCTTTGGAAAAGTTGCCAAGAAATAGAGCGCGGCGAGTACATTCCTTTCTATCAACAA 647
Db 6 CCCATTAGAAAAGCTTTGGAGAGCTAGAGGCAAGATGGTATATCTGTCTACCAAGAT 65
Qy 648 CAGATTTCTCATACAAAGACTTTTAACTTCTAGTTAGTTAGTTCAATTTGCTTCTAG 707
Db 66 GAAGCTTTCCATGATATAAATCTTTACTAGAGCTAGGAAATTAGATTTCAATCTAGTGCAG 125
Qy 708 TCATTGCAAGGAAGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAG 767
Db 126 TCATGCAAGGAAGAGCTTAGCAATCTTCCAGAGTGGTGGAAAGAAATAGACTTTGCT 185
Qy 768 AAGAAGCACTTGTGTTTAAAGATAGAAATTTGAAATGCTACTTTTGGGAGTACGTTTCA 827
Db 186 ACAAGTTACTTTTGCACGAGACAGATTGGTTCAAGGCTACTTTTGGATGATGCGGCTG 245
Qy 828 GGTATGAGCAGACAGTATTCGCGGCTAGAGTTTCTTCAAAAGAGCTTTGCTGTATA 887
Db 246 TATTTTGGCCCCCAATCTTACGGGGTAGACGAATTTCAACAAAGTAATGCCATGACA 305
Qy 888 ACTCTTATAGATGACACTTATGATGCGTATGCTTATGAAGAACTTAAAGATCTTTTACT 947
Db 306 TCCATTTCTAGATGATATCCATGATGATGCGCACCTCGAAGACTCAAGCTCTTCATA 365
Qy 948 GAAGCTTTGAAAAGTGGTCAATTTACATGCTTTAGACACACTTCCAGATACATGAACCG 1007
Db 366 GAAGCCATTGAGAGATGGGATATTAAACAGCATAAATCAGCTTCCAGAAATACATGAACCTC 425
Qy 1008 ATATACAAATTTTATGATGATACATACACAGAAATGGAGAAATTTCTTTCRAAGGAGGA 1067
Db 426 TGCTATGTCGCACCTCTTAGATGTGTACAAAGAAATCGAGGAAGAGATGGAGAAAGAGA 485
Qy 1068 AGAAGATCTTATTTAACTGCGGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATG 1127
Db 486 AACCAATTCGGTTTCACTATGCCAAGAAAGTAATGAAGAAATCAAGTTCCAGCTTACTTT 545
Qy 1128 GTTGAACCAAAATCGGCAAAATGAGGAGACATACCAACCATGAAGAGATGATCCAGTT 1187
Db 546 GCCGAGGCCAAATGGTTACATGAAGAAACACGTCAGCAATTTGAAGAGTACATGCGTGT 605
Qy 1188 GTAAT 1192
Db 606 GCACT 610

RESULT 8
CD486420
LOCUS
779 bp mRNA linear EST 01-JUL-2003
CD486420

DEFINITION CRH1.4G10 Cotton Root and Hypocotyl Lambda ZIPLOX Library (CRH)
Gossypium hirsutum cDNA clone CRH1.4G10 5', similar to
(+)-delta-cadinene synthase, mRNA sequence.
CD486420
CD486420.1 GI:31407395
EST.
Gossypium hirsutum (upland cotton)
Gossypium hirsutum
Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 779)
Dowd,C., Wilson,I. and McFadden,H.
Different Gene Expression Responses in Cotton Root and Hypocotyl
tissues during infection with Fusarium Wilt Disease
Unpublished (2003)
Contact: Caitriona Dowd, Helen McFadden
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industry
Black Mountain Laboratories, Cnr Clunies Ross Street & Barry Drive,
Black Mountain, Canberra, ACT, 2601, Australia
Tel: 61 2 6246 4914, 6246 5377
Fax: 61 2 6246 5000
Email: Caitriona.Dowd@csiro.au, Helen.McFadden@csiro.au
Vector clipped sequences Bases 1-17 (GTCGACCCACGGCTCG): SalI
adapter
Seq primer: M13 reverse primer
High quality sequence stop: 779.
Location/Qualifiers
1. .779
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="DeltaEMERALD"
/db_xref="taxon:3635"
/clone="CRH1.4G10"
/tissue_type="Root and hypocotyl tissues"
/dev_stage="5 day old seedlings"
/lab_host="Y1090(ZL)"
/clone_lib="Cotton Root and Hypocotyl Lambda ZIPLOX
Library (CRH)"
/note="vector: Lambda ZIPLOX; Site 1: SalI; Site 2: NotI;
mRNA was prepared from root and hypocotyl tissues of the
cotton cultivar DeltaEMERALD. cDNA was synthesised from a
NotI-oligo dT primer/adaptor using the manufacturers
protocols (Life Technologies) and then ligated to a SalI
adaptor to facilitate directional cloning. The cDNA was
cloned into the SalI and NotI sites of the Lambda ZIPLOX
phage vector (Life Technologies). Constructed by Caitriona
Dowd and Helen McFadden."

ORIGIN
Query Match 14.8%; Score 244; DB 14; Length 779;
Best Local Similarity 59.3%; Pred. No. 1.2e-50;
Matches 456; Conservative 0; Mismatches 301; Indels 12; Gaps 2;
QY 228 CCGTATCATTGACGGGAGATGATCATCGATTCGATTTATTAAGAACATATCGT 287
DB 19 CCGGTTCATTTCACGAGGAGTCGAGATGAACTAGAGATATCTCATTACACCAAT 78
QY 288 GATAACTGGAATG---GTGACCGCTCTTCCTTATGTTTGGTTCGTTATGCGAAAGCAAGGA 344
DB 79 GATGCGGAGAACGACCTCTACACCATCCCTTCGATTCGACTACTCCGAGAGCATGGA 138
QY 345 TATTATGTTACATGATGATGTTTCAATAACTATAAAGCAAAATGAGCGCTTCAGCAA 404
DB 139 TTCAATGTTTCATGCGACGATTTCAACAAAGTTTAAAGACGAGGGAATTTCAAGTCA 198
QY 405 TCCTTACTAATGATGTTGAAGGTTTGTCTGAGTTGTACGAGCAACTTCTATGAGGTA 464
DB 199 TCGTGTACAGCGATGTTGAGGATGTTGGAACCTTACCAAGCTTCTATTTGAGGTT 258
QY 465 CCTGGGAGATTTATTAGAGATGCTCTTGTTTTTACACGATCTCGCTTTAGCATTTATG 524

Db 259 CATGGGGAAGATATATTGGATGAAGCAATTTCTTTCCACGACCAACCATTTAAAGCCTT--- 315
QY 525 ACNAAAGATGCTTTTCTACAAACCCCGCTCTTTTACCGAAATACAAACGGGACCTAAAG 584
Db 316 -----GCAGTAGCATCTTTGGACCATCTTTATCCGAAGAGGTTTCTCATGCTTTGAAA 369
QY 585 CAACCCCTTTGGAAAAGGTTGGCAAGATAGAGGGCGCGGAGTACATCTCTTTCTATCAA 644
Db 370 CAATCAATTCGAAGAGGCTTGGCAAGGTTGAGGCAAGACACTATCTTTTCAGTATACCAA 429
QY 645 CAACAAGATTTCTCATACAAGACTTTTACTTAAGTTGCTAAAGTTAGAGTTCAATTTGCTT 704
Db 430 GATATTGAGTCCCATTAATTAAGGTTTGTGGAGTTTGTCTAAGATCGATTTCAACATGGTA 489
QY 705 CAGTCATTGCAACAAGGAGAGCTCAGCCATGTGTGCAAAATGTGTGAAAGCTTTTCGATATC 764
Db 490 CAACCTTTTGCATAGAAAAGAGCTAAGTGAGATTCTTAGGTGTGGAGGATTTAGACTTT 549
QY 765 AAGAAGACGCACTTGTGTTTAAGAGATAGATAATGTTGAAATGCTACTTTTGGGACTAGGT 824
Db 550 CAAGAAGATTGCCATATACGCAAGAGATAGATGTTGTAAGGCTATTTTTCGATCTCAGGA 609
QY 825 TCAGGCTATGAGCCACAGATATTCCCGGCTAGAGTTTCTTCACAAAGCTGTTGCTGTT 884
Db 610 GTGTACTTTGAGCCCCCAATATTCTCTTGTGTAGAAAGATGTTGACAAAGTATAGCAATG 669
QY 885 ATAACTCTTATAGATGACACTTATGATGCTATGATGCTATGATGCTATGATGCTATGATGCT 944
Db 670 GCATCTATTGTAGATGATACATATGACTCATATGCAACATATGAGAGCTCATTCCTTAT 729
QY 945 ACTGAAGCTGTTGAAAGGTTGGTCAATTATCATGCTTTAGACACACTTCCAG 993
Db 730 ACAATGCAATTGAGAGGTGGGATATCAAAATGATAGATGAACTTCTGCT 778

RESULT 9
AY105371
LOCUS AY105371 1981 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO100777 mRNA sequence.
ACCESSION AY105371
VERSION AY105371.1 GI:21208449
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1981)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 1981)
AUTHORS Coe,E.H.
JOURNAL Direct Submission
REFERENCE Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
TITLE If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR,
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
LOCATION/Qualifiers
1. .1981
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635786"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"

/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 14.7%; Score 243.2; DB 11; Length 1981;
Best Local Similarity 50.5%; Pred No. 2.3e-50;
Matches 739; Conservative 0; Mismatches 698; Indels 27; Gaps 5;

QY 38 TGCCAACTTTCTCCCAAGCAATTTGGGAGATGAGTTTCTCTATCAATCAAAAGCAAGTAGA 97
DB 113 TCCCTCTCTCCATCCAACTCTGTGGGAGATTTCTTTCTTTCTTACGAGGCGCCAGCGA 172
QY 98 GCAAGGGGTGAACAGATAGTGAATTTAAAGAGAGTGGCGCACTACTAAAGA 157
DB 173 G---GCACAGAAAGCTGAAATGCGCAAGGGCTGAAGTGTGGGAGAGAGTGAAGA 229
QY 158 AGCTTTGGATATTCCTATGAACATGCCAAATTTGTTGAAGCTGATTGATGAAATTCACG 217
DB 230 CATGATTAAGGTTCAATGATGTGCCAGAAATAGTGACCTTATATCACCTGCAGCG 289
QY 218 CTTTGAATACCGTATCATCTTTGAACGGGAGATTGATCATGCAATGCAATGATTTATGA 277
DB 290 ACTAAATCTGACTACCACTACGAGGATGAGATAAATGAAAGTTGGCTGTGTTTACAA 349
QY 278 AACATATGGTGATTAACCTGAATGTGACCGCTCTTCCTATGTTGGTTCGCTTATCGAAA 337
DB 350 CTCGAAATATGATGGTGTAAATTTGGACTTATGTTTTCACGCGGATTTCTATCTTCGCTAA 409
QY 338 GCAAGGATATTTATGTTACATGTGATTTTCAATAACTATATAAGACAAAATGGAGCGTT 397
DB 410 ATGTGGCTATCATGTTTCATCGATGATTTCTAAATTTTCAAGACCAATATGGAAATTT 469
QY 398 CAAGCAATCGTTAGTAAATGATTTGAAGTTTGTCTTGAATGTTACGAGCAACTTCTAT 457
DB 470 TA-----TTGAGGTAGATACCGAGGCTTGTAAAGCTTATACAAATGACAGCATACCT 520
QY 458 GAGGTACTCTGGGAGATATATTAGAGATGCTCTTGGTTTATACAGATCTCTGCTTAG 517
DB 521 GAGATTCATCGAGACAGACAGTCTGTGATGAGCAATTTCTTTTATACGAGATGCTTCA 580
QY 518 CATATGACAAAAGATGCTTTTCTACAAAACCCGCTCTTTTATCCGAAATACACGGGC 577
DB 581 AGATAGATTAGAACATTTGGATCACCAATAGCAG-----AGGAGTCTGCTGTC 631
QY 578 ACTAAGCAACCCCTTTGGAAAAGTTGCCAAGAAATAGAGCGCGCGCAAGTACATTCCTTT 637
DB 632 CTTTGATACCCCACTTTTGAAGGGTTGGAACTTTGGAAATGAAGGATTACATTCCTAT 691
QY 638 CTATCAACAAAGATTTCTCATACAAAGCTTACTTAACTTGTCTAAGTTAGATTCAA 697
DB 692 CTACGAAAGGACGCTTAAACAAAACAAATCAATATTTAGAGTTTGGAAACTGAATCTCAA 751
QY 698 TTTGCTTCAGTCATTGCAACAGGAAGAGCTCAGCCATGTGCAATATGTTGGAAGCTTT 757
DB 752 CCTTCTACAGCTTTCTTATTCATCAGAGCTCAAGAGTGCACAACTGTGTGAAGAATCT 811
QY 758 CGATATCAAGAAAGACCGCACTTTGTTTAAGAGATAGAAATGTTGAATGCTACTTTTGGGG 817
DB 812 CCGCGCTGAATCAAACTTGAGTTTGTGAGGGATAGATAGTAGAGGTGCTACTTTGGAT 871
QY 818 ACTAGGTTCAAGCTATGAGCCACAGTATTCGCGGCTAGAGTTTCTTCCAAAAGCTGT 877
DB 872 GAGTGGGGGATGCTATGATCTCTCAATTTCTCATTTCTCGAATATATATTTACAAAGATTGT 931
QY 878 TGCTGTTTAACTCTTATAGATGACATATCATGCTGATGTTGTTACTTATGAAGACTTAA 937
DB 932 TGCCTTCATCAGATCTAGATGATACCCCTTCACTCAGATGCTAATTTCTTATGAGGAT 991
QY 938 GATCTTTACTGAAGCTGTGAAAGGTGTCAATTATCATGCTTTAGACACACTTCCAGAATA 997

DB 992 GCAACTTGCAGAGCAGTCGAAAGGTGGGATGAAAGTGCAGTCTCTACTATCCAGATA 1051
QY 998 CATGAAACCGATATACAAATTTATTCATGGATACATACAGAAATGGAAGAAATTTCTTGC 1057
DB 1052 CATGAAGGATTTCTACATGATTTGTTGAAGACATTTTCATCGTTTGAAGATGATTAGG 1111
QY 1058 AAAGAGGGGAAGAACAGATCTATTTAACTGCGCAAGAAATTTGTGAAAGATTTGTTAG 1117
DB 1112 ACCAGA---CAAGAGCTACCGAGTGTCTTACTTTAAAGAAAGCGTGAACGGAATTAGTTCG 1168
QY 1118 AACCTGATGGTTGAAGCAAAATGGCCAAATGAGGACACATACC---AACCACTGAAGA 1174
DB 1169 GGAATATACAAAGAAATAAATGGCGTGATGAGGATTTATGTCCTCCAAAACATTTGAAGA 1228
QY 1175 GCATGATCCAGTTGTTAATCACTTACTGCGGCTGCTAACTGCTTACAAACAATTTGTTATCT 1234
DB 1229 ACACATAAAGTTTCACTCATAGCATTTGAGGTACACTGCTGCTGTGTTCCGCGTTTGT 1288
QY 1235 TGGCATGATGATATATTCACAAAAGAGTCTGCGAATGGGCTGTCTCTGCACTCTCTCT 1294
DB 1289 TGGAAATGGGTGATGTAGTAACAAAGAAAGATCATGGAATGGGTTATGATGATGCTGAACT 1348
QY 1295 TTTTATGATCTCAGGTATATCTTGGTCGAGCGCTTAATGATCTCATGCCCAACAGCCGA 1354
DB 1349 TGTTAAGTCTTTTGGTATATTTGTACGGCTCTCCAATGATTTGATCAACCAAGCGTGA 1408
QY 1355 GCAAGAAAGAAACATAGTTTCATCGAGCGCTTGAAGTTATATGAAGAAATATAATGTCAA 1414
DB 1409 ACAAGAGAGAGCATTTGTCTCCACTGTTTACATGAGCAGCATGAACATAAC 1468
QY 1415 TGAGGATGATGCCAAACCTTTGATTTTCAAGGAAGTAGAAGATGTTGGAAAGATATAAA 1474
DB 1469 AATGGATGAAGCATGTGAGCAGATAAAGAACTTACAGAAAGATTCATGGAAGTTTCATGAT 1528
QY 1475 CCGAGAGTACCTCACAACCTAAAAA 1498
DB 1529 TGAACAGGCGCTTGCACATAAAGA 1552

RESULT 10

AU293055 628 bp mRNA linear EST 04-DEC-2002
AU293055 zinnia cultured mesophyll cell equalized cDNA zinnia
LOCUS
DEFINITION
elegans cDNA clone Z7794, mRNA sequence.

ACCESSION
AU293055

VERSION
AU293055.1 GI:24253563

KEYWORDS
EST.

SOURCE
Zinnia elegans

ORGANISM
Zinnia elegans

Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M.,
Matsuo, N., Minami, A., Nagata-Hiwatashi, M., Nakamura, K.,
Okamura, Y., Sassa, N., Suzuki, S., Yazaki, J., Kikuchi, S. and
Fukuda, H.

REFERENCE

AUTHORS

TITLE

Visualization by comprehensive microarray analysis of gene
expression programs during transdifferentiation of mesophyll cells
into xylem cells

JOURNAL

COMMENT

Proc. Natl. Acad. Sci. U.S.A. 99 (24), 15794-15799 (2002)

Contact: Taku Demura

Morphogenesis Research Group

RIKEN Plant Science Center

1-7-22 Suehirocho, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9605

Fax: 81-45-503-9573

Email: demura@postman.riken.go.jp

This clone was obtained at our laboratory.

Seq primer: M13 forward.

Location/Qualifiers

```
source          1..628
/organism="Zinnia elegans"
/mol_type="mRNA"
/cultivar="Canary bird"
/db_xref="taxon:34245"
/clone="Z7794"
/tissue_type="mesophyll cell"
/clone_lib="zinnia cultured mesophyll cell equalized cDNA"
/notes="vector: pGEM-T easy; cultured in tracheary element
differentiation-inductive medium"

ORIGIN
Query Match          14.5%; Score 239.8; DB 9; Length 628;
Best Local Similarity 64.1%; Pred. No. 1.3e-49;
Matches 361; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 591 CTTTGAAGAGGTTGCCAAGATAGAGCGCGCAGTACATCTCTTCTATCAACAA 650
Db      |||||
QY 1 CTTGAAAAGATTTCCAGGCTAGAGCGTTGGTTACATGCCAATCTCAACAA 60
Db      |||||
QY 651 GATTCTCATCAAGACTTTACTTTAACTGCTAAGTTAGAGTTCAATTTGCTTCAGTCA 710
Db      |||||
QY 61 GAATCTCATATCATGACTTGTCTAACACTTGTCTAAGTTAGATTTCAATCTACTCCAGAT 120
Db      |||||
QY 711 TTGCACAGAGAGCTCAGCCATGTGTGCAATGTGTGGAAGCTTCGATATCAAGAG 770
Db      |||||
QY 121 TTGCACGAAAGGAGCTTAGCCGAATTAGCAAGGGTGGAAAGAGCTTGTATCTCAAA 180
Db      |||||
QY 771 AACGCACCTTGTTTAAGAGATAGATTTGTAATGCTACTTTTGGGGACTAGGTTTCAGGC 830
Db      |||||
QY 181 AACTACTTTATGTTCCGGACAGAAATAGTCGAAGTTACTTTTGGATATTGGCGGTCTAC 240
Db      |||||
QY 831 TATGAGCCACAGTATTCCTCCGGCTAGAGTTTCTTCAAAAAGCTTGTGCTTTATACT 890
Db      |||||
QY 241 TTCGAGCTCAACATTCGAATCAAGGATTTTCTTGAAGGTTGCAATCTAGTGATC 300
Db      |||||
QY 891 CTTATAGATCACACTTATGATGCTATGATGTTACTTTATGAGAACTTAAGATCTTACTGAA 950
Db      |||||
QY 301 ATTTGGATGACACATATGATACTATGTTGTTTATGAAGAACTCGAGATCTTTACAGAA 360
Db      |||||
QY 951 GCTGTTGAAAGGTTGCTCAATTACATGCTTACACACACTTCCAGAAATACATCAAAACCGATA 1010
Db      |||||
QY 361 GCAAGTTCAAAATGTTGTAAGTTGATGATGATGCTTCTCTGATATCATGAACCGATA 420
Db      |||||
QY 1011 TACAAATATTATCGATACATACACAGAAATGGAAGAATTTCTTGCAAAAGGAGGA 1070
Db      |||||
QY 421 TACCAAGAGCTTCTCGATGTGCACAGAGCCGAATGTGTACTTTGAGAGACAGAAAG 480
Db      |||||
QY 1071 ACAGATCTATTAACTCCGGCAAGAAATTTGTGAAGAGTTTGTAGAACTGATGTT 1130
Db      |||||
QY 481 ACATACGTTATCTATCATACGAGATATGTTTAAAGCGTACACTCGAACTTTCTGAGTC 540
Db      |||||
QY 1131 GAAGCAAAATGGGCAATGAGGG 1153
Db      |||||
QY 541 GAGCCGAATGGCNAAGAGGG 563
Db      |||||

RESULT 11
LOCUS      BG645114
DEFINITION EST506733 KV3 Medicago truncatula cDNA clone pkV3-39C12 5' end,
mRNA sequence.
ACCESSION BG645114
VERSION   BG645114.1 GI:13780226
KEYWORDS  EST.
SOURCE    Medicago truncatula (barrel medic)
ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE 1 (bases 1 to 762)
AUTHORS   VandenBosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D.,

Bowman,C.L., Craven,M.B., Cho,J. and Fraser,C.M.
ESTs from roots of Medicago truncatula 72 h after Rhizobium
inoculation, 2001
Unpublished (2001)
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cbs.umn.edu
M393521e TIGR sequence name: MTEBX18TK More information is
available at: www.medicago.org
Seq primer: SKmod (CTA GAA CTA gtg gat CC).
Location/Qualifiers
1..762
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pkV3-39C12"
/tissue_type="Seedling roots"
/dev_stage="3 days post-inoculation with Sinorhizobium
meliloti"
/lab_host="E. coli strain XL0LR"
/clone_lib="KV3"
/notes="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."
```

Db 527 TAGATCTTTACAAAGAAATAGAGCAGGAGATGAAAAAGAGGAAGAGAAATATGCACCTTA 586
 QY 1084 ACTCGGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGG 1143
 Db 587 ACTACTACGTGAAGAAATTCATAAATACGTTCAAGCATATATAGTCTGAGGCCAGATGCT 646
 QY 1144 CAATATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTCTGTAATCAATTACTGGGG 1203
 Db 647 TAAATGATAATATCAACCAACATAGAGAGTACATTCGCATATCAACAGAAATCATGTG 706
 QY 1204 GTGCTAACCTCTTACAAACACTTGTATCTTGGCATGAGTGATATATTCACAAA 1259
 Db 707 GTTACGCACTGTTACAAACACTTGTATCTTCAATTTTCATGGGGATACAGCTACAGAA 762

RESULT 12
 CB968730 753 bp mRNA linear EST 30-APR-2003
 LOCUS CAB10002 Iva Fa_F06 Cabernet Sauvignon Flower Pre-bloom - CAB1
 DEFINITION Vitis vinifera cDNA clone CAB10002_Iva_Fa_F06 5', mRNA sequence.
 ACCESSION CB968730
 VERSION CB968730.1 GI:30252091
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 753)
 Goes da Silva, F., Randolino, A., Lim, H., Baek, J., Jones, K. and
 Cook, D.
 Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 CONTACT Contact: Douglas Cook, PhD
 CABS Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcocok@ucdavis.edu
 Seq primer: ACGGTACCGGCATATGCC.
 Location/Qualifiers
 1..753
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultiivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB10002 Iva Fa_F06"
 /sex="Hermaphrodite"
 /dev_stage="Pre-bloom"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"
 /note="Organ: Flower - Pre-bloom; Vector: pDNR; Site 1:
 Sfil; Site 2: Sfil; CAB1 is a cDNA library of Vitis
 vinifera cv. 'Cabernet Sauvignon' clone 8 flowers. Samples
 were collected approximately eleven days before onset of
 bloom (clusters at this stage were fully developed and
 flowers with calypters or caps still attached. Sampled
 vines were located at the University of California, Davis,
 Experimental Vineyard. cDNAs were made by oligo-dT priming
 and directionally cloned. 5' and 3' adaptors were used in
 cloning as follows:
 5'-AAGCAGTGGTATCAACGAGAGTGGCCATTACGGCGGG-3' and
 5'-ATTCTAGAGCGGAGCGGCGGACATG-dt(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 14.3%; Score 235.6; DB 14; Length 753;
 Best Local Similarity 57.8%; Pred. No. 1.5e-48;
 Matches 435; Conservative 0; Mismatches 316; Indels 1; Gaps 1;

QY 889 CTCCTATAGATGACACTTATGTCGCTATGCTTATGAAGAACTTTAAGATCTTTACTG 948
 Db 2 CTAATTTATGATGACATCTATGACGCTATGCTTACCTTTGAAGAACTTCGAGCTCTTCACCG 61
 QY 949 AAGCTGTTGAAGAGTGGTCAATTTACATGCTTAGACACACTTCCAGAATACATGAACCGA 1008
 Db 62 AAGCAGTTGAGAGTGGGATGTCACTGCCATAGATCAACTGCCAGAGTACATGAGAGTGT 121
 QY 1009 TATACAAATTTATCATGGATACATACACAGAAATGGAAGAAATTTCTTGAAGAGGGAA 1068
 Db 122 GTTATCAAGCACTCTTATATGTATACAGTGAATTTGAGGAAGAAATGCCCAAGGAAGAA 181
 QY 1069 GAACAGATCTATTAACTGCGGCAAGAAATTTCTGAAAGAGTTTGTAGAAACCTGTATGG 1128
 Db 182 GATCGTATCGGCTCTACTATGCAAAAGAGCAATGAAATCAAGTAAGGACACACTATG 241
 QY 1129 TTGAAGCAAAATGGGCAAAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTG 1188
 Db 242 AGGAAGCCAAATGGCTTCAAGTCAACAAATTCACCACTGAGGAGTACATGCTGTGG 301
 QY 1189 TAATCATTTACTGGCGTGTAACTCTTACACAACTTCTTATCTTGGCATGAGTGATA 1248
 Db 302 CGTTAGTTACCTCTGCATCTCAATGCTGGCGACCATCGTTTGTGGAAATGGAGATG 361
 QY 1249 TATTCAAAAAGAGTCTGCGAAATGGGCTGCTCTGCACTCTCTCTTTTAGATACTCAG 1308
 Db 362 CTGTGACAAAAGAAACCTTCGATTGGATTTTCAGTGAGCCTAAGATTGTAAGGGCTTCAG 421
 QY 1309 GTATACTTGGTGACGCGCTAAATGATCTCATGCCCAAGCGGCGGACGAGAAGAAAC 1369
 Db 422 CCATAGTTTGCAGGCTCATGGATGATGTTGTTTCCCAAAAGTTTGAGCAGAAGAGAGAC 481
 QY 1369 ATAGTTTCATCGAGCCTTGAAGTTATATGAAGAAATATATGTCAATGAGGAGTATGCC 1428
 Db 482 ATGTTGCCCTCAGCTGTTGAATGTTATATGAAGCAACATGTCGCTCGAGCAGGAAGCGC 541
 QY 1429 AAACCTTGTATTTACAAAGAGTAGAAGATGTGTGGAAGATATAAAACCGAGAGTACCTCA 1488
 Db 542 ACAACGAATTTAAACAACAAGTAAGAGATGATGATGAAGAGTATCAATGAAGAGTGCCTCA 601
 QY 1489 CAACATAAAACATTCOAAGCGCTTATGATGCTGTGATCTATTTGTGCCAGTTCTCTG 1548
 Db 602 TCCCAACAGCTGTCCCAATGCCCATCTCATGGAGTTCTTAAATCTCGCAGCTGTTATAG 661
 QY 1549 AAGTTCAATA-TGCAGGAAGGATACTTCACACGATGGGAGACGAATACAAACATCTC 1607
 Db 662 NTGTTATATACAGAATGAAGATGGGCTATCTCATCTGNACTGTTCTTAAGGATTTTC 721
 QY 1608 ATAAAGTCTCTACTCGTTTATCTCTATGAGTAT 1639
 Db 722 GTAACCTCCATGCTAATGATCTCTGTGCCAAT 753

RESULT 13
 AW685590
 LOCUS NF029D03NRI.F1000 Nodulated root Medicago truncatula cDNA clone
 DEFINITION NF029D03NR 5', mRNA sequence.
 ACCESSION AW685590
 VERSION AW685590.1 GI:7560326
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eucosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 663)
 Watson, B.S., Shin, H.S., Lopez-Meyer, M., Scott, A.D., Harris, A.R.,
 Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W.,
 May, G.D. and Paiva, N.L.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula nodulated root library

JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Paiva NL
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7317
 Fax: 580 221 7380
 Email: nlpaiva@noble.org
 Insert Length: 663 Std Error: 0.00
 Plate: 029 row: D column: 03
 Seq primer: TCACACAGGAACAGCTATGAC.
 Location/Qualifiers
 1. .663
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF029D03NR"
 /tissue_type="root"
 /dev_stage="Pooled developmental"
 /clone_lib="Modulated root"
 /note="Vector: Lambda Zap; Four-week old Rhizobium
 mellotii-inoculated Medicago truncatula roots, containing
 a mixture of young and old roots and nodules."

ORIGIN

Query Match 14.1%; Score 232.6; DB 10; Length 663;
Best Local Similarity 60.0%; Pred. No. 8.5e-48;
Matches 388; Conservative 0; Mismatches 259; Indels 0; Gaps 0

629	Qy	CATTCTTTCTATCAACAACAGAAATTCTCAATAACAAGACTTTTACTATAAACTTGTAAAGTT	688
1	Db	CATTCTCTATGAGAAAGATCTCTACCTGTGATGAAGTTCCTACTCACTTTTCGCAAAAT	60
689	Qy	AGAGTTCAAATTTGCTTCAGTCAATGCCAAGGAGAGCTCAGCCATGTCGCAATGCTG	748
61	Db	GGATTCAAATTTACTCCAAATCCCTACATCAAAAGAGTTTGAAACATTTCCAAATGCTG	120
749	Qy	GAAAGCTTTTCGATATCAAGAAAGACGCACTTGTTTAAGAGATAGAAATGTTGAATGCTA	808
121	Db	GAAGAGTTGGACTTCTCTACCAAACTACCTTATGCAAGAGATAGGATCGCGGAATGTAG	180
809	Qy	CTTTTGGGCTAGGTCAGGCTATGAGCCACAGATATTCGCCGGCTAGAGTTTCTTCAC	868
181	Db	CTTTTGGGTTTGACTGCAATTTTGAACCCCAATATTTCTCAAGCAAGAAATGATGAT	240
869	Qy	AAAAGCTTTGCTGTATACTCTATAGATGACACTTATCATGGTATGTGTACTTATGA	928
241	Db	TAAAGTAATCACCCCTACTTTTCAATCATGTGATGATACATATGATGCATATGAACTTGA	300
929	Qy	AGAACTTAAGATCTTTACTGAAGCTGTTGAAAGTGTCATATCATGCTTTAGACACACT	988
301	Db	TGAAGTGGAACTTTTATCCAAAGCGAGTTGAAAGGTGGGATATTAGCAGCTTTGGACGA	360
989	Qy	TCCAGAAATACATGAACCGATATACAAATTTATCATGATACATACACAGAAATGGAGA	1048
361	Db	TCCAGACTCATGAAACCTTCTATAGATCATCTTTAAACAATTTACGAGAAATAGAGAA	420
1049	Qy	ATTTCTTCGAAGGAGGGAAGAACAGATCTATTTAACTCGCGCAAGAAATTTGTGAAGA	1108
421	Db	AGAAATCAGAAAAAGAGGAGGATATATACCCCTTGACTACTACAAAAATGGAATTCAAAA	480
1109	Qy	GTTTGTTTAGAAACCTGATGTTGAAGCAAAATGGGCATAATGAGGAGCATACCAACCA	1168
481	Db	GTCAGTCCAAAGCATTTATGACAGAGGCGAGATGGCTGAATGAAACCCATATACCAACCA	540
1169	Qy	TGAAGAGCATGATCCAGCTGTGTAATCATTTACTTGGCGGTGCTAACTGCTTACAAACCTT	1228
541	Db	AGAAGAGTACATGGCATATCAAAAAATCAGGTGCTTACCCCTTGTGATATTAACTTC	600
1229	Qy	TTATCTTGGCATGAGTGATATATTCAAAAAAGAGTCTGTGAAATGGG	1275
601	Db	TTACATTTGATGGGTGACATAGCCACAAAGAGGATCTTCAATTTGGG	647

ORIGIN

	Query Match	14.0%;	Score 230.6;	DB 14;	Length 777;
	Best Local Similarity	58.3%;	Pred. No. 2.8e-47;		
	Matches 404;	Conservative 0;	Mismatches 289;	Indels 0;	Gaps 0;
Qy	705	CAGTCATTCACAAGGAAGAGCTCAGCCATGTGTGCAAAATGTTGGAAAGCTTTTCGATATC	764		
Db	29	CAATTTTTCATAGGAAAGAGCTAGCGAGATTCACGCTGTTGGAAAGATTTAGACTTT	88		
Qy	765	AAGAAGAAGCCACTTGTTTTAAAGATAGAAATTTGTGAAATGCTACTTTTGGGACTAGGT	824		
Db	89	CAAGAAGAGTTGCCATATCCACAGAGATAGAGTGGTTGAGGCTATTTTGGATCTCAGGA	148		
Qy	825	TCAGGCTATGAGCCAGTATTCCCGGGCTAGAGTTTTCTTCACAAAAGCTGTGCTGTT	884		

RESULT 14	
CD485858	
LOCUS	
DEFINITION	

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE AUTHORS TITLE	JOURNAL COMMENT
<p>1. <i>Journal of the American Medical Association</i>, 1964, 191: 1000-1001.</p> <p>2. <i>Journal of the American Medical Association</i>, 1964, 191: 1000-1001.</p>	<p>1. <i>Journal of the American Medical Association</i>, 1964, 191: 1000-1001.</p> <p>2. <i>Journal of the American Medical Association</i>, 1964, 191: 1000-1001.</p>

FEATURES
SOURCE

CD485858 777 bp mRNA linear EST 01-JUL-2003
 CFUS3.4F02 Cotton Root and Hypocotyl tissues infected with Fusarium
 oxysporum f.sp. vasinfectum (Fov) Lambda ZIPLOX Library (CFUS)
 Gossypium hirsutum cDNA clone CFUS3.4F02 5' similar to
 (+)-delta-cadinene synthase, mRNA sequence.
 CD485858
 CD485858.1 GI:31406823
 EST.
 Gossypium hirsutum (upland cotton)
 Gossypium hirsutum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 777)
 Dowd, C., Wilson, I. and McFadden, H.
 Different Gene Expression Responses in Cotton Root and Hypocotyl
 tissues during infection with Fusarium Wilt Disease
 Unpublished (2003)
 Contact: Caitriona Dowd, Helen McFadden
 Commonwealth Scientific and Industrial
 Research Organisation
 Division of Plant Industry
 Black Mountain Laboratories, Cnr Clunies Ross Street & Barry Drive,
 Black Mountain, Canberra, ACT, 2601, Australia
 Tel: 61 2 6246 4914, 6246 5377
 Fax: 61 2 6246 5000
 Email: Caitriona.Dowd@csiro.au, Helen.McFadden@csiro.au
 Vector clipped sequences Bases 1-17 (GTGACCCACGCGTCCG): SalI
 adapter
 Seq primer: M13 reverse primer
 High quality sequence stop: 777.

Db 149 GTGACTTTGAGCCCAATATTCACCTGGTAGAAGATGTTGACAAAGTAGCAATG 208
Qy 885 ATAACTCTTATAGACACTTATGATGGTATGGTACCTTATGAAGAACTTAAGATCTTT 944
Db 209 GCATCTATTGATGATACATATGACTCATATGCAACATATGAAGAGCTCATTCATAT 268
Qy 945 ACTGAAGCTGTTGAAGGTTGCAATATACATGCTTACACACACTTCCAGAAATACATGAA 1004
Db 269 ACAAATGCAATGAGAGGTTGGGATATCAATGATAGTGAATCTTCTGTAATACATGAG 328
Qy 1005 CCGATATACAAATATTATTCATGATACATACACAGAAATGGAAGAAATTTCTTGCAAGGAG 1064
Db 329 CCGAGCTACAAGCACTATTAGATGTTTACGAGAATGGAACAACTGGTGCCTGAGCAT 388
Qy 1065 GGAAGAACAGATCTATTAACTCGCGCAAGAAATTTGTGGAAGAGTTTGTGAAGACCTG 1124
Db 389 GGAAGACATATCGTGCATATGCAAAATATGCAAAATATGCGATGATACGACTTGTCTAATCTTAC 448
Qy 1125 ATGGTTGAAGCAAAATGGCAAAATGAGGACACATACCAACCACTGAAGAGCATGATCCA 1184
Db 449 CTGTGGAGCCAGATGGACTCTTCAAACTACAACCATCATTCGAGGAGTTTAGGCT 508
Qy 1185 GTTGAATCATTAAGTGGCGTGTAACTGCTTACAACTGTTTATCTTGGCATGAGT 1244
Db 509 AATGCAATGCCAACTTGTGTTATGCAATGCTTGTATTACATCTTTCGTCGCAATGGA 568
Qy 1245 GATATATTCACAAAGAGTCTGCGAATGGGCTGTCTCTGACCTCTCTTTTATGATAC 1304
Db 569 GATATCGTAACCAAGAGACTTTAAATGGCGACCAATGACCTTAAGATCATTCAGCT 628
Qy 1305 TCAGGTATTAAGTGTGCGAGCGCTTAATGATCTCATGACCAACAGGCGGAGCAAGAA 1364
Db 629 TCACAATTAATTTAGTGTATGATGATGTTGCGGAACACAAAGTTTAAGCATAGGAGA 688
Qy 1365 AAACATAGTTTCATGCGAGCGCTTGAAGTTATATG 1397
Db 689 GAAGACGATGCTCAGCAATAGATGTTACATG 721

RESULT 15
BG588994
LOCUS
DEFINITION BG588994 807 bp mRNA linear EST 12-APR-2001
EST490803 MHRP- Medicago truncatula cDNA clone pMHRP-59114, mRNA
sequence.
ACCESSION BG588994
VERSION BG588994.1 GI:13607134
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 807)
AUTHORS Harrison,M.J.; Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
and Fraser,C.M.
TITLES ESTs from phosphate-starved roots of Medicago truncatula, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-221-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
The Samuel Roberts Noble Foundation: N387745e TIGR sequence name:
MTHB555K More information is available at: http://www.medicago.org
Seq primer: SKmd (CTA GAA GTG GAT CC).
Location/Qualifiers
FEATURES
source
1..807
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"

/db_xref="taxon:3880"
/clones="pMHRP-59114"
/tissue_type="roots"
/dev_stage="phosphate-starved"
/lab_host="XLOR"
/clone_lib="MHRP-"
/note=Vector: pBlueScript SK-; Site.1: EcoRI; Site.2:
XhoI; At the trifoliolate stage, M. truncatula plants were
transplanted to phosphate-free sand and grown for a
further 30 days. During this period, they were fertilized
twice weekly with 1/2 Hoaglands solutions containing 20mM
potassium phosphate. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into the
Unizap XR vector from Stratagene and packaged using
GibcoPack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-zap
phage using Ex-assist helper phage and propagated in
XLOR cells."

ORIGIN

Query Match 13.7%; Score 226.4; DB 12; Length 807;
Best Local Similarity 56.8%; Pred. No. 3.3e-46;
Matches 436; Conservative 0; Mismatches 331; Indels 1; Gaps 1;
Qy 883 TTATAACTTTATAGATGACACTTATGATGCTATGCTATGCTTATGAAGAACTTAAGATCT 942
Db 3 TAATAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 62
Qy 943 TTACTGAAGCTGTTGAAGGTTGTCATTTACATGCTTTAGACACACTTCCAGAAATACATGA 1002
Db 63 TTACTAAGCACTCGAAAGGTTGGGATATTAGTTGCTTGGATAATCTTCCGGATTACATGA 122
Qy 1003 AACCGATATACAAATTTATCATGATATACATACACAGAAATGGAAGAAATTTCTTGCAAGG 1062
Db 123 AGTTTCTCTACGTGATATCTTAGATCTTTACAAAGAAATAGAGCAGGAGATGAAAGAAAG 182
Qy 1063 AGGGAAGAACAGATCTATTAACTCGCGCAAGAAATTTGTGAAAGAGTTGTTAGAAACC 1122
Db 183 AAGGAGAGAAATATGCACTTAATCTAGTGAAGAAATTCATAATATGATGATGATGATGAT 242
Qy 1123 TGATGTTTGAAGCAAAATGCGCAATAGAGGACACATACCAACCACTGAAGAGCATGATC 1182
Db 243 ATATGACTGAGCCAGATGTTAAATGATATAATATCAACCAACATTTAGAAAGATGATATTC 302
Qy 1183 CAGTTGTAATCATTTACTGCGGTTGCTTAACCTGCTTACACAACTGTTTATCTTGGATGA 1242
Db 303 GCATATCAACAGAAATCATGTTGTTACGCACTGTTTACACAACTGTTTATCTTTCATGG 362
Qy 1243 GTGATATATTCACAAAGAGTCTGTGCAATGGGCTGTCTCTGCACTCTCTCTTTTATAGAT 1302
Db 363 GGGATACAGCTACAGAGACATCTTCAGTGGGTATCAATATGGGCCAAATAATGATCAATG 422
Qy 1303 ACTCAGGTATATCTTGGTCAAGCCCTAAATGATCTCATGACCCCAAGAGCGGAGCAAGAAA 1362
Db 423 CTGCTATTATCTTTGAGGCTTAATGGATGATTTGTCATCCCAAGAGTTTGAACAAAAAA 482
Qy 1363 GAAACATAGTTTCATCGAGCCCTTGAAGTTATATGAGGAATATATATGTCATGAGGAGT 1422
Db 483 GAGATCATGTTTCTCTCATCTTTGSAATGTTATATGAGCAACATAATATCTCTAGAGAAG 542
Qy 1423 ATGCCCAACCTTGTATTTACAGGAAGTAGAAGATGTTGTTGATGCTTGGAAAGATATGAAAT 1482
Db 543 GTGCCATTCAGAGAGTTCGAAAAGAAATTTGTTGATGCTTGGAAAGATATGAAATAGGAAT 602
Qy 1483 ACCTCACAACATAAAACATTTCCAGGCCCTTATGATGGCTGTGATCTATTTTGTGCGAGT 1542
Db 603 GTCTTATCCGACTGAAATTTCCATCCCTTTTGTGACGTGCACTTTAAACATTTATCACGTT 662
Qy 1543 TTCTTGAAGTTCAATATGCGAGAAAGGATACTTACACGTATGGGAGAGATACAAAC 1602
Db 663 TCATGGATGTGTTTATAAAGATAAAGATACTTACACATCCCTGAGGAGAAATGAAGA 722
Qy 1603 ATCTCATAAAGTCTCT-ACTCGTTTATCTCTATGATGATGAGGATCC 1649

Db 723 CATTCAATCAATCTTAACTTGTGATCCAGTGCCCAATTGAGCATAC 770

Search completed: June 7, 2004, 12:38:38
Job time : 4510 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 16:37:26 ; Search time 45 Seconds
(without alignments)
3849.322 Million cell updates/sec

Title: US-09-763-822A-14

Perfect score: 2906

Sequence: 1 TMLTEKPIRPIANFPPI.....GDEYKHLIKLLVPMISGS 549

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	2888	99.4	546	10 Q9FY41	Q9fy41 artemia a
2	2885	99.3	546	10 Q9FVMS	Q9fvm5 artemia a
3	2878	99.0	546	10 Q9LW98	Q9lww8 artemia a
4	2864	98.6	546	10 Q9AR04	Q9ar04 artemia a
5	1730	59.5	548	10 Q9SA63	Q9sa63 artemia a
6	1699.5	58.5	549	10 Q9FXV6	Q9fxv6 artemia a
7	1577.5	54.3	549	10 Q9AR67	Q9ar67 solidago ca
8	1463.5	50.4	547	10 Q9ST45	Q9st45 artemia a
9	1462.5	50.3	547	10 Q9LLR9	Q9lllr9 artemia a
10	1314.5	45.2	573	10 Q9AXP5	Q9axp5 artemia a
11	1303.5	44.9	577	10 Q9FXV7	Q9fxv7 artemia a
12	1183	40.7	544	10 Q9FQ27	Q9fq27 lycopersico
13	1156	39.8	546	10 Q9FQ26	Q9fq26 lycopersico
14	1143	39.3	555	10 Q9SANO	Q9san0 gossypium a
15	1139.5	39.2	555	10 Q9FYU6	Q9fyu6 citrus juno
16	1135	39.1	548	10 Q64961	Q64961 lycopersico

17	1130	38.9	551	10 Q9LKN1	Q9lkn1 gossypium h
18	1130	38.9	552	10 Q9FQ28	Q9fq28 lycopersico
19	1123	38.6	548	10 Q64962	Q64962 lycopersico
20	1092	37.6	548	10 Q8RVR2	Q8rvr2 citrus para
21	1063.5	36.6	551	10 Q9ZTQ6	Q9ztq6 solanum tub
22	1060	36.5	550	10 Q9SBJ0	Q9sbj0 solanum tub
23	1059	36.4	556	10 Q9ZTQ7	Q9ztq7 solanum tub
24	1056	36.3	555	10 Q84LK9	Q84lk9 gossypium a
25	1055	36.3	549	10 Q9XJ25	Q9xj25 solanum tub
26	1041	35.8	556	10 Q9XJ32	Q9xj32 solanum tub
27	1040.5	35.8	557	10 Q9ZTQ8	Q9ztq8 solanum tub
28	1039	35.8	549	10 Q9XIZ0	Q9xiz0 solanum tub
29	1038	35.7	598	10 Q9FVL3	Q9fvl3 nicotiana t
30	1037.5	35.7	556	10 Q9FVL3	Q9fvl3 lycopersico
31	1036.5	35.7	559	10 Q81923	Q81923 capsicum an
32	1032.5	35.5	559	10 Q65323	Q65323 capsicum an
33	1028.5	35.4	553	10 Q9ATN6	Q9atn6 capsicum an
34	1020	35.1	560	10 Q9SDN9	Q9sdn9 capsicum an
35	1017	35.0	597	10 Q941H1	Q941h1 nicotiana t
36	1012.5	34.8	548	10 Q84LF0	Q84lf0 nicotiana a
37	1009.5	34.7	548	10 Q84LF2	Q84lf2 nicotiana a
38	1005.5	34.6	561	10 Q81634	Q81634 elaeis olei
39	999.5	34.4	560	10 Q94J58	Q94j58 citrus juno
40	996.5	34.3	548	10 Q84LGO	Q84lgo nicotiana a
41	992	34.1	508	10 Q9SW77	Q9sw77 gossypium a
42	988.5	34.0	558	10 Q8LSC2	Q8lsc2 cichorium i
43	984.5	33.9	559	10 Q8S3A6	Q8s3a6 lactuca sat
44	983.5	33.8	548	10 Q84LF1	Q84lf1 nicotiana a
45	978	33.7	520	10 Q39978	Q39978 hyoscyamus

ALIGNMENTS

RESULT 1

Q9FY41 ID Q9FY41 PRELIMINARY; PRT; 546 AA.
AC Q9FY41;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Amorpha-4,11-diene synthase.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteraceae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21183247; PubMed=11289612;
RA Wallaart T.E., Bouwmeester H.J., Hille J., Poppinga L., Majers N.C.;
RT "Amorpha-4,11-diene synthase: cloning and functional expression of a
RT key enzyme in the biosynthetic pathway of the novel antimalarial drug
RT artemisinin";
RL Planta 212:460-465(2001).
DR EMBL; AY006482; AAF98444.1; --
DR HSPB; Q40577; SEAU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cycToirid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 546 AA; 63932 MW; FCD786A7ED81FB21 CRC64;

Query Match 99.4%; Score 2888; DB 10; Length 546;

Best Local Similarity 99.8%; Pred.No. 3.le-206;

Matches 545; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MALTEKPIRPIANFPPIWGDSQFLYKQVEQVEQVINDLKKEVRQLKEALDIPMKH 61

Db 1 MSLTEKPIRPIANFPFSGDQFLIYKQVQGVQVIVNDLKKEVRQLLKEALDIPMKH 60
QY 62 ANLLKLIDEIQRGLGIPYHFEREIDHALQCIYETVYGNWNGDRSSLWFLMRKOGYVTTCD 121
Db 61 ANLLKLIDEIQRGLGIPYHFEREIDHALQCIYETVYGNWNGDRSSLWFLMRKOGYVTTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 181
Db 121 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 180
QY 182 TNPALFTEIQRALKQPLWKLRIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 241
Db 181 TNPALFTEIQRALKQPLWKLRIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 240
QY 242 ELSHVCKWKAFAFIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 301
Db 241 ELSHVCKWKAFAFIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGTYEELKIPTFAVERWSITCLDTLPYMKPIYKLFMDTYTMEBEFLAKEGRDILF 361
Db 301 TYDAYGTYEELKIPTFAVERWSITCLDTLPYMKPIYKLFMDTYTMEBEFLAKEGRDILF 360
QY 482 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 541
Db 481 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 540
QY 542 VYPMSI 547
Db 541 VYPMSI 546

RESULT 2

Q9FM5 PRELIMINARY; PRT; 546 AA.
ID Q9FM5 PRELIMINARY; PRT; 546 AA.
AC Q9FM5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amorpha-4,11-diene synthase.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=20485153; PubMed=11032404;
RA Merck P.E., Bengtsson M., Boumeester H.J., Posthumus M.A.,
RA Brodelius P.E.;
RT "Molecular cloning, expression, and characterization of a amorpho-
4,11-diene synthase from, a key enzyme of artemisinin biosynthesis of
Artemisia annua L.";
RL Arch. Biochem. Biophys. 381:173-180(2000).
DR EMBL; AF138959; RA61439.1; --
DR HSSP; Q40577; SEAU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp-cyc.oxid.
DR InterPro; IPR001906; Terp-synth-like.
DR Pfam; PF01397; Terpene_synth; 1.

DR Pfam; PF03936; Terpene_synth_C; 1.
SQ SEQUENCE 546 AA; 63933 MW; 381386A7ED81FEF9 CRC64;
Query Match 99.3%; Score 2885; DB 10; Length 546;
Best Local Similarity 99.6%; Pred. No. 5.1e-206;
Matches 544; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 MALTEEKPIRPIANFPFSGDQFLIYKQVQGVQVIVNDLKKEVRQLLKEALDIPMKH 61
Db 1 MSLTEKPIRPIANFPFSGDQFLIYKQVQGVQVIVNDLKKEVRQLLKEALDIPMKH 60
QY 62 ANLLKLIDEIQRGLGIPYHFEREIDHALQCIYETVYGNWNGDRSSLWFLMRKOGYVTTCD 121
Db 61 ANLLKLIDEIQRGLGIPYHFEREIDHALQCIYETVYGNWNGDRSSLWFLMRKOGYVTTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 181
Db 121 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 180
QY 182 TNPALFTEIQRALKQPLWKLRIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 241
Db 181 TNPALFTEIQRALKQPLWKLRIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 240
QY 242 ELSHVCKWKAFAFIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 301
Db 241 ELSHVCKWKAFAFIKQAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGTYEELKIPTFAVERWSITCLDTLPYMKPIYKLFMDTYTMEBEFLAKEGRDILF 361
Db 301 TYDAYGTYEELKIPTFAVERWSITCLDTLPYMKPIYKLFMDTYTMEBEFLAKEGRDILF 360
QY 362 NCGKEFVKFVFNLMVAKWANEHGIPTTEHDPPVITGGANLLTTTCYLGMSDIFTKE 421
Db 361 NCGKEFVKFVFNLMVAKWANEHGIPTTEHDPPVITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVWAVSAPPLFRYSGLGRRLNDLMTKABQERKHSSSLESYMKYNNVNEEYATLIY 481
Db 421 SVWAVSAPPLFRYSGLGRRLNDLMTKABQERKHSSSLESYMKYNNVNEEYATLIY 480
QY 482 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 541
Db 481 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 540
QY 542 VYPMSI 547
Db 541 VYPMSI 546

RESULT 3

Q9LM98 PRELIMINARY; PRT; 546 AA.
ID Q9LM98 PRELIMINARY; PRT; 546 AA.
AC Q9LM98;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amorpha-4,11-diene synthase.
GN KCS12.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21036108; PubMed=11185551;
RA Chang Y.U., Song S.H., Park S.H., Kim S.U.;
RT "Amorpha-4,11-diene synthase of Artemisia annua: cDNA isolation and
RT bacterial expression of a terpene synthase involved in artemisinin
RT biosynthesis.";
RL Arch. Biochem. Biophys. 383:178-184(2000).
DR EMBL; AJ251751; CAB94691.1; --
DR HSSP; Q40577; SEAT.

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Beta-caryophyllene synthase QHS1.
 GN QHS1.
 OS Artemisia annua (Sweet wormwood).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Asterales; Asteraceae; Asteroidae; Anthemideae;
 OC Artemisia.
 OX NCBI_TaxID=35608;
 RN SEQUENCE FROM N.A.
 RP Cai Y., Crock J., Jia J.-W., Lin Z.-X., Croteau R., Chen X.-Y.;
 RT "Beta-Caryophyllene Synthase From Artemisia annua L.: cDNA Isolation,
 RL Characterization, Wound Induction and Elicitation.";
 DR EMBL; AF472361; AAL79181.1; -;
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR005630; Terpene synthase C.
 DR InterPro; IPR008949; Terpenoid synth.
 DR InterPro; IPR008930; Terp_cyc_foroid.
 DR Pfam; PF01397; Terpene synth-like.
 DR Pfam; PF03936; Terpene synth; 1.
 DR SEQUENCE 548 AA; 63737 MW; 40B8E69751FF6654 CRC64;
 Query Match 59.5%; Score 1730; DB 10; Length 548;
 Best Local Similarity 58.0%; Pred. No. 3.9e-120;
 Matches 318; Conservative 101; Mismatches 127; Indels 2; Gaps 2;
 QY 2 MALTTEKTRTPANPPSPWGDQFLIY-QKQVEQ-GVEQIVNDLKKEVRLKEALDIPM 59
 DB 1 MSVKEEKVIRPIVHFPPSWADQFLIDDKQAEQANVEQVNVNELRDVRLDVLSSLDVQT 60
 QY 60 KHANLLKLDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLPFLMRKQGYVT 119
 DB 61 EHTNLLKLDAIQRGLIAYHFEETEQALQHIYDTYGDWNGKSPSLPFLMRKQGYVYS 120
 QY 120 CDVFNKYKDKNGAFKQSLANDVEGLLELYEATSRVPEGEIILDLGFTSRSLSMTKDA 179
 DB 121 CDI FKNYKEDSGFRESLNDVEGLLELYEATYLRVQGEGLDLDALVFTTCLEKIAKDL 180
 QY 180 FSTNPALFTEIQRLAKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLH 239
 DB 181 VHTNPTLSYIQEALQKPLKRLTRLEALRYIPMYEQQASHNESLLAKLGFNLLQSLH 240
 QY 240 KEELSHVCKWKAFTDKKNAFCPLDRIVECYFGLGSGYEPQYSRVRPFTKAVAVITLI 299
 DB 241 RKELSEVSRWKGDLVFNPLPYARDRMVECYFVALGVYFEPKYSQARIFLAKVISLATVL 300
 QY 300 DDTYDAYGYEELKIFTEAVERWSITCLDTPLEVMKPIYKLFMDTYTEMEEFLEKEGRTD 359
 DB 301 DDTYDAYGYEELKIFTEAVERWSITCLDTPLEVMKPIYKLFMDTYTEMEEFLEKEGRTD 360
 QY 360 LNCQKGFVEFVKNLMVEAKWANEHGIPTTEHDPVVIITGGANLLTTTCYLGMSDIFT 419
 DB 361 HLSYAKESKKEFTIRSYMMEAKWANEHGVPTAEHMSVAFVSSGYSMLATTCFVGMGDIVT 420
 QY 420 KESVWAVSAPPLFRYSGLGRNLDMTHKABOERKHSSSLSYKMEYNNVEEVAQTL 479
 DB 421 DEAFKWTLPPIIKASCAIARLMDDDIHSQKEEKIRHVASSVESYNNQYDVTTEHVLKV 480
 QY 480 IYKEVEDVWKDINREYLTNNIPRLLMAVILYQCFLEVOYAGKONFTRMGDHYKHLIKS 539
 DB 481 FNKKIEDAWKDIRESLVRKDI FVPLMLRVINLAQVMDVLYKHKGDTNVGEEELKDHIKS 540
 QY 540 LLVYPMWI 547
 DB 541 LLVHPPI 548

RESULT 6
 Q9FYX6

Q9FYX6 PRELIMINARY; PRT; 549 AA.
 AC Q9FYX6;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative sesquiterpene cyclase.
 GN CASC34.
 OS Artemisia annua (Sweet wormwood).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Asterales; Asteraceae; Asteroidae; Anthemideae;
 OC Artemisia.
 OX NCBI_TaxID=35608;
 RN SEQUENCE FROM N.A.
 RP TISSUE=Leaf;
 RC PubMed=10996256;
 RA Van Geldre E.G., De Pauw I., Inze D., Van Montagu M.,
 RT "Cloning and molecular analysis of two new sesquiterpene cyclases from
 RL Artemisia annua L.";
 DR Plant Sci. 158:163-171(2000).
 DR EMBL; AJ271793; CAC12732.1; -;
 DR HSPF; Q40577; SEAT.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR005630; Terpene synthase C.
 DR InterPro; IPR008949; Terpenoid synth.
 DR InterPro; IPR008930; Terp_cyc_foroid.
 DR InterPro; IPR001906; Terp_synth-like.
 DR Pfam; PF01397; Terpene synth; 1.
 DR Pfam; PF03936; Terpene synth; 1.
 DR SEQUENCE 549 AA; 63808 MW; DADBE5C8244BACFE CRC64;
 Query Match 58.5%; Score 1699.5; DB 10; Length 549;
 Best Local Similarity 57.2%; Pred. No. 7.3e-118;
 Matches 314; Conservative 101; Mismatches 131; Indels 3; Gaps 3;
 QY 2 MALTTEKTRTPANPPSPWGDQFLIY-QKQVEQ-GVEQIVNDLKKEVRLKEALDIPM 59
 DB 1 MSVKEEKVIRPIVHFPPSWADQFLIDDKQAEQANVEQVNVNELRDVRLDVLSSLDVQT 60
 QY 60 KHANLLKLDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLPFLMRKQGYVT 119
 DB 61 EHTNLLKLDAIQRGLIAYHFEETEQALQHIYDTYGDWNGKSPSLPFLMRKQGYVYS 120
 QY 120 CDVFNKYKDKNGA-FKQSLANDVEGLLELYEATSRVPEGEIILDLGFTSRSLSMTKD 178
 DB 121 CDI FKNYKEDSGFRESLNDVEGLLELYEATYLRVQGEGLDLDALVFTTCLEKIAKDL 180
 QY 179 AFSTNPALFTEIQRLAKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSL 238
 DB 181 LVHTNPTLSYIQEALQKPLKRLARLEALHYIPMYEQQASHNESLLAKLGFNLLQSL 240
 QY 239 HKELSHVCKWKAFTDKKNAFCPLDRIVECYFGLGSGYEPQYSRVRPFTKAVAVITL 298
 DB 241 HRKELSEVSRWKGDLVFNPLPYARDRMVECYFVALGVYFEPKYSQARIFLAKVISLATV 300
 QY 299 IDDTYDAYGYEELKIFTEAVERWSITCLDTPLEVMKPIYKLFMDTYTEMEEFLEKEGR 358
 DB 301 LDDTYDAYGYEELKIFTEAVERWSITCLDTPLEVMKPIYKLFMDTYTEMEEFLEKEGR 360
 QY 359 DLFCNGKGFVEFVKNLMVEAKWANEHGIPTTEHDPVVIITGGANLLTTTCYLGMSDIF 418
 DB 361 HLSYAKESKKEFTIRSYMMEAKWANEHGVPTAEHMSVAFVSSGYSMLATTCFVGMGDIV 420
 QY 419 KESVWAVSAPPLFRYSGLGRNLDMTHKABOERKHSSSLSYKMEYNNVEEVAQTL 478
 DB 421 DEAFKWTLPPIIKASCAIARLMDDDIHSQKEEKIRHVASSVESYNNQYDVTTEHVLK 480
 QY 479 LIYKEVEDVWKDINREYLTNNIPRLLMAVILYQCFLEVOYAGKONFTRMGDHYKHLIK 538
 DB 481 VFNKKIEDAWKDIRESLVRKDI FVPLMLRVINLAQVMDVLYKHKGDTNVGEEELKDHIK 540

Db 32 VRNTVKPPPSWIGDQFLTYHPPKSLTIEKHQVEV-----EIRKKLLNTAFSEPKQHT 86
QY 63 NLLKLEIDRIGLPIYHFEREIDHALOCITYTYGDNWNG-----DRSSLWFLMRKQGYV 118
Db 87 KLHLHDSVORLGVSHFEQIEBALQHVYATHGQWIGKDNLXSTQWFLRQCGENV 146
QY 119 TCDVFNKYDKNGAFKOSLANDVEGLELEATSRVPGEIILEDALGFTSRSLIMTYD 178
Db 147 SSGIFENHMDKGFKEHLSDDVQGLGLYEAAVMSVEGEKELDDALEFTKHLGNIAD 206
QY 179 APTNPALEIQRALKQPLWKRIPRIEAAQYIPFYQOQDSHNKTLKLALENNLQSL 238
Db 207 P-SQNASLRTKIBQALNRPLRKMRLEALHYIPYQOEASHDETLLNLAKLDYNNMLOSI 265
QY 239 HKEELSHVCKWKAFADIKKNAPCLDRIVECYFWGLSGYBPQYSRARVFTTKAVAVITL 298
Db 266 HKEEISEICKWKWDLSFNKLPVDRDLVEIYFWLGVYFPQHYRSMFLTKTSMWLV 325
QY 299 IDTYDAYGYEELKIITEAVERWSITCLDPLPEYMKPIYKLFMDTYTEMEEFKAGRT 358
Db 326 LDDTYDNYGYEELKIITEAVERWSITCLDPLPEYMKPIYKLFMDTYTEMEEFKAGRT 385
QY 359 DLFNCGEFKEFVRNLWAKWANEHGIPTTEEDHPVVIITGGANLLTTCYLGMSDIF 418
Db 386 YQFYAKEMVKECRNLLVEAKWLKEGVPTLEBHSVSCVYAYAVMIANSYVGRNNV 445
QY 419 TKESVEMAVSAPPLFRYSGLGRRLNDLMTKABQERKSHSSLESYMKYNNVEEYAT 478
Db 446 TESFKWAKVFPPLVATCLILRLMDLADIAGEEERDHHVVSIECYRNETCASEDAVK 505
QY 479 LIYKEVEDVWKDINREYLTTRKIPRLIMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIK 538
Db 506 FLKSQVEDAKWKINKECLRPTEIPNSLTIPPLSLARVSDILYKTNNGYNHAGEEIVSIYK 565
QY 539 SLAVYPM 545
Db 566 SLIVHPL 572

RESULT 11

Q9FYX7 PRELIMINARY; PRT; 577 AA.
ID Q9FYX7
AC Q9FYX7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sesquiterpene cyclase.
GN CASC125.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC Artemisia.
CX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX PubMed=10996256;
RA Van Geldre E.G.; De Pauw I., Inze D., Van Montagu M.,
RA Van den Beekhout E.;
RT "Cloning and molecular analysis of two new sesquiterpene cyclases from
RT Artemisia annua L."
RL Plant Sci. 158:163-171(2000).
DR EMEL; AJ271792; CAC12731.1; -.
DR HSP; Q4577; SEAT.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.

QY 5 TEEKIRPIANPPPSWIGDQFLTYHPPKSLTIEKHQVEV-----EIRKKLLNTAFSEPKQHT 86
Db 26 TQPDVIRHTMNFNASWIGDQFLTYHPPKSLTIEKHQVEV-----EIRKKLLNTAFSEPKQHT 86
QY 59 MKHANLLKIDRIGLPIYHFEREIDHALOCITYTYGDNWNG-----NGDRSSLWFLMRKQ 114
Db 83 MOHVKLIELIDAVQRIGIAHYHFEREIDHALOCITYTYGDNWNG-----NGDRSSLWFLMRKQ 142
QY 115 GYVTCDFVNNKDKNGAFKOSLANDVEGLELEATSRVPGEIILEDALGFTSRSLIMTYD 178
Db 143 GNVSSGVFKDFMDEKGEKESLNCDAQGLALYEAFAFMEVEDETLIDNALEFTKHLGNIAD 202
QY 175 MYKDAFSTNPALFTEIQRALKQPLWKRIPRIEAAQYIPFYQOQDSHNKTLKLALENNLQSL 234
Db 203 IAKDP-SCDSSLRTQIHALKQPLRRLRLARIEALHYPIYQOQDSHNKTLKLALENNLQSL 261
QY 235 LQSLHKEELSHVCKWKAFADIKKNAPCLDRIVECYFWGLSGYBPQYSRARVFTTKAVAVITL 294
Db 262 LOSMEKKEELSHVCKWKAFADIKKNAPCLDRIVECYFWGLSGYBPQYSRARVFTTKAVAVITL 321
QY 295 VITLIDDTYDAYGYEELKIITEAVERWSITCLDPLPEYMKPIYKLFMDTYTEMEEFKAGRT 354
Db 322 MLVLLDDTDFNYGYEELKIITEAVERWSITCLDPLPEYMKPIYKLFMDTYTEMEEFKAGRT 381
QY 355 BORTDLFN--CKGEFKEFVRNLWAKWANEHGIPTTEEDHPVVIITGGANLLTTCYLGMSDIF 410
Db 382 GKG-NISNSLCQGRWKQELGSLTIVETKAKGVHQAQPLEEYMSVSNVGTGTYGLMIAARS 440
QY 411 YLGMSDIFTKESVEMAVSAPPLFRYSGLGRRLNDLMTKABQERKSHSSLESYMKYNNVEEYAT 470
Db 441 YVGRGDIVTDFKVVSSYPPPIKASCIVIRLMDLADIAGEEERDHHVVSIECYRNETCASEDAVK 500
QY 471 VNEEYATLIYKEVEDVWKDINREYLTTRKIPRLIMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIK 530
Db 501 ASEEEACEYISRKVEDAKWKINREYLTTRKIPRLIMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIK 560
QY 531 DEYKHLIKSLIVYPM 547
Db 561 GDMKSYMKSFVHPMVV 577

RESULT 12

Q9FQ27 PRELIMINARY; PRT; 544 AA.
ID Q9FQ27
AC Q9FQ27
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sesquiterpene synthase 1.
GN SSTLH1.
OS Lycopersicon hirsutum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
CX NCBI_TaxID=62890;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GV. LA1777;
RX MEDLINE=20541624; PubMed=11090225;
RA van Der Hoeven R.S., Monforte A.J., Breeden D., Tanksley S.D.,
RA Steffens J.C.;
RT "Genetic Control and Evolution of Sesquiterpene Biosynthesis in
RT Lycopersicon esculentum and Lycopersicon hirsutum."
RL Plant Cell 12:2283-2294(2000).
DR EMEL; AF279455; AAG41891.1; -.
DR HSP; Q40577; SEAT.
DR GO; GO:0016829; F:lyase activity; IEA.

```

DR GO:0008152; P:metabolism; IEA.
DR InterPro: IPR005630; Terpene synth. C.
DR InterPro: IPR008949; Terpenoid synth.
DR InterPro: IPR008930; Terp. cyc. tocid.
DR InterPro: IPR001906; Terp. synth.-like.
DR Pfam: PF01397; Terpene synth; 1.
DR Pfam: PF03936; Terpene synth_C; 1.
DR SQ SEQUENCE 544 AA; 63723 MW; 16EA14FCDEA048BB CRC64;

Query Match 40.7%; Score 1183; DB 10; Length 544;
Best Local Similarity 41.3%; Pred. No. 1.8e-79;
Matches 226; Conservative 118; Mismatches 199; Indels 4; Gaps 2

QY 2 MALTEEPRIPIANFPFSGINGDQLIYKQVQGVQEVNDLKKEVRQLLEALDIPMKH 61
DB 1 MAASSANKSREPLANFHTVWCYHFLSYTHEITNQEKVEVDEYKETIRKQVLEA---PGS 57

QY 62 ANLLKLDEIQRIGIPHYHFEREIHQAQCIYETYGDNWNGDRSSLWFLMRKQGYVTC 121
DB 58 EQKLVLIDAMQRLGVAYHFHNEIETSIQNI FDPAKQNNNLHIVLSRFLVRVQGHYMSD 117

QY 122 VFNNYKDKNGAFKQSLANDVIGLLELYEATSWRPVGEIILEDALGFTSRSLISMTKDAPS 181
DB 118 VFQOFTNQDGKFKEITLNDVQGLLSLYEASVLRVRDEEILEALFAFTTHLKSIVSTWSN 177

QY 182 TNPALFTIQRALKQPLWKRLPRIEAAQIIPYQOQDSHNKTLKLAKLEENLLQSLHKE 241
DB 178 NNNSLKVESEALTQPTIMTLPRWEARYISYIYENNDAHLLLKPAKLDNFNNLQKLQR 237

QY 242 ELSHVCWKWAKFADIKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVITLDD 301
DB 238 ELSDLTRWKDLDLPANKYPYARDRLVECYFWILGVYPEPKYSRARKQMTKVLKMTSIIDD 297

QY 302 TYDAYGYEBELKIITEAVERNSITCLDTLPYMKPIYKLPMDTYTEMEEFIAKSGRTDLF 361
DB 298 TFDAYATFDELEPFNDAIQRWDANAIDSIPPYMRPAYQAFLDIYSEMEQVLSKXGKLDRV 357

QY 362 NCGKEFYKEFVRLNLMVEAKWANE-CHPTTBEHDPVVIITGGANLLTTTCVLGMSDIFTK 420
DB 358 YYAKNVEKVLVRAYFKETQWLNDCHDHPKYBEHMENSLVSGYGMIPPTCLVGMEEFISI 417

QY 421 ESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSHSSLESITKMYENNYNEEVAQTLI 480
DB 418 ETTEWLMDPLIVPASSLIARAMNDIVGHEVEQERGHVASLIECMKDYGASKQEAQAKF 477

QY 481 YKEVEDYWKDINREYLTATTONIPRELLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 540
DB 478 KKVQVTNAWKDINKKEFFRTEVPFMFVLRLVNLTRAABEPLYKEKDAYTNAGKGLKNMINSI 537

QY 541 LVTPMSI 547
DB 538 LIESVKI 544

```

```

RESULT 13
Q9FQ26
ID Q9FQ26 PRELIMINARY; PRT; 546 AA.
AC Q9FQ26;
DT 01-MAR-2001 (TREMBlurel. 16, Created)
DT 01-MAR-2001 (TREMBlurel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlurel. 25, Last annotation update)
DE Sesquiterpene synthase 2.
GN SSSLH2.
OS Lycopersicon hirsutum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamnids; Solanales; Solanaceae; Solanum.
CC NCBI_TaxID=62890;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. LA1777;
RC MEDLINE=20541624; PubMed=11090225;
RA van Der Hoeven R.S.; Monforte A.J.; Bredsen D.; Tankaley S.D.;

```

RA	Steffens J.C.;
Rt	"Genetic Control and Evolution of Sesquiterpene Biosynthesis in
Rr	Lycopersicon esculentum and Lycopersicon hirsutum.";
RL	Plant Cell 12:2283-2294(2000).
DR	ENBL; AF279456; ARG41892.1; -.
DR	HSP; Q40577; SEAU.
DR	GO: GO:0016829; F:lyase activity; IEA.
DR	GO: GO:0008152; P:metabolism; IEA.
DR	InterPro: IPR005630; Terpene synth. C.
DR	InterPro: IPR008949; Terpenoid synth.
DR	InterPro: IPR008930; Terp_cyc_boroid.
DR	InterPro: IPR001908; Terp_synth-like.
DR	Pfam: PF01397; Terpene synth.; I.
DR	Pfam: PF03936; Terpene synth. C; 1.
SQ	SEQUENCE 546 AA; 64045 MW; 19C49B2264BBCA3D CRC64;

Query Match	39.8%; Score 1156; DB 10; Length 546;
Best Local Similarity	40.5%; Pred.No. 1.9e-77;
Matches 223; Conservative 117; Mismatches 202; Indels 8; Gaps 4;	

Qy	2	MALTEEKPIRPIANFPFSGIWGDOLFIYKQVEGVEQIVNDLKKVRLKLKEALDIPMKH	61
Dd	1	MAASFANKCRPLANPHTPVGVGHFLYPNBITNQEKIEVDYEKETIRKMLVEA--PEGS	57
Qy	62	ANLLKLIDEIQRIGIPYHFEREIDHALOCIYETYGDNWNGDR---SSLWPRLMRKQYYV	118
Dd	58	EQLVLVIDAMORLGVAHYFHFEHNETSTQNIFDAPKQN-NDDNLHIYSLRFLVRQGHYM	116
Qy	119	TCDFVNNYKDKNAGAFQSGLANDVEGLELEYEATSMVPGBIILEDALGFTRSLSTMTKD	178
Dd	117	SSDVFFQFTNQDGKFTELTDVQGULLSYEAHLRVNRNEELDEALTFTPTTHLESIVSN	176
Qy	179	AFTSNPALFTEIORALKQPLWKLPRIEAAQYTFPYQQODSHNKTLTLLKLAKLEFNLIQSL	238
Dd	177	LSNKNSLKVEVSSEALSQPIMTLPRIGAERYISYENDAHNHLLJKFAKLDNFMLQKP	236
Qy	239	HKSELGHVCWKWKAFDIKKNAPCLRDRIVECTFWGLGSGYEPGYSPARVFTFKAVAVITL	298
Dd	237	HQRELSDLTRWMDLDPANKIPYARDRLVECYFWILGVTFPKYSARKMMTKVLKMTSI	296
Qy	299	IDDTYDAYGTYYBELKIFTAVERNSITCLDTLPVMKPIYKLPMDTYTWEEEFIAKEGR	358
Dd	297	IDDTFYAYANFDELVPFNDAIQBDANAIDSIPPYMRPIYQALLDIYGEWDQVLSKEGKL	356
Qy	359	DLFCNGKEFKVEFRVNLVWEAKWAN-EGHPTTEEHDPPVIITGGANLLTTTCYLGMDSI	417
Dd	357	DRVYAKYENKLVRAYFKESOWLNDNDHPKYEEHEHNAIVTVGYVMGATNCLVGMEEF	416
Qy	418	FTKESVEMAVSAPPLFRYSGILGRRLNDLWTHKAEQERKHSSSSLSYKEYNVNBEYAQ	477
Dd	417	ISKTEFWLMSEBPVIVRASSLSIGRAMDDIVGHEVEQERGHVASIIECYMKDYGASKQEAY	476
Qy	478	TLYIKEYVEDVKNDINREYLTTKNIAPPLMAVILYCQFLEVQYAGKNDFTRMDEVKHLI	537
Dd	477	AKPKVDVTNAWKDINKEFFRTPTEVPFMVLERALNFTRTIDTLYQEBEDVYNAGKLUKMI	536
Qy	538	KSLLVYPMISI 547	
Dd	537	HSILIPS VKI 546	

RESULT 14	
Q9SANO	
ID	Q9SANO PRELIMINARY; PRS; 555 AA.
AC	Q9SANO;
DT	01-MAY-2000 (TEMBlrel. 13, Created)
DT	01-MAY-2000 (TEMBlrel. 13, Last sequence update)
DT	01-OCT-2003 (TEMBlrel. 25, Last annotation update)
DE	(-)delta-cadinene synthase.
GN	CATI-A.
OS	Gossypium arboreum (Tree cotton).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

RESULT 14

O9SANO

ID Q9SANO

AC Q9SANO

DT 01-MAY-

DT 01-MAY-

DT 01-OCT-

DE (+) - del

GN CAT1-A.

OS
Gossyp

OC Eukaryo

OC
Spermat

OC	eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX	NCBI_TaxID=29729;
RN	SEQUENCE FROM N.A.
RP	STRAIN=cv. Nanjing.
RC	Liang W.Q., Tan X.P., Heinstein P., Yamada Y., Hashimoto T.,
RA	Chen X.Y.
RT	"expression pattern of a (+)-delta-cadinene synthase gene CADI-A and
RT	sequiterpene aldehyde biosynthesis in seedlings of Gossypium arboreum
RT	L.";
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; Y18484; CAA77191.1; -.
DR	HSP; Q40577; SEAT.
DR	GO: GO:0016829; F-lyase activity; IEA.
DR	GO: GO:0008152; P-metabolism; IEA.
DR	GO: GO:0008152; P-metabolism; IEA.
DR	InterPro; IPR005630; Terpene synth. C.
DR	InterPro; IPR008949; Terpenoid synth.
DR	InterPro; IPR008930; Terp. cyc. toroid.
DR	InterPro; IPR001906; Terp. synth-like.
DR	Pfam; PF01397; Terpene synth. C; 1.
DR	Pfam; PF03936; Terpene synth. C; 1.
SQ	SEQUENCE 555 AA; 64104 MW; 70846394B133D53B CRC64;
Query Match	
Best Local Similarity 39.3%; Score 1143; DB 10; Length 555;	
Matches 226; Conservative 111; Mismatches 197; Indels 18; Gaps 6;	
QY	3 ALTEEKPIRIANFPPIGWQDILY-QKQVEQGVQVNDLKEVRLKALDIPMKH 61
DB	15 AISEEN--RPKADPHGIGWDMFIICDPTDIDAATELQYBELKAQVKRMINEFVD---DS 69
QY	62 ANLLKALIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGD-----RSSLWFLMRKQG 115
DB	70 NQKLPRFIDAVQRLGVSVYFEKEIEDELENIYR--DTNNNDADTLYTALRFLRLREHG 126
QY	116 YVTCDFVNNYKXNGAFKQSLANDVEGLELEAYEATMRVPGIILEDALGPFTRSLIM 175
DB	127 FDISCAFNFKDEAGNFKASLTSDVQGLLELYEASMYRHGEDILDEAISFTTAQLTLA 186
QY	176 TKDAFSTNPALFTIQRALKQPLWKRLPRTEAAQYIPFYQQDSHNKTLKLAKLEFNLL 235
DB	187 LP---TLHPLSQVGHALKQSTIRGLPRVEARNFTSIYQDLSHNKSLLOPAKIDENLL 243
QY	236 QSLHKBELSHVCKWKAFTDKNAPCLDRDRIVECYFWGLSGYEPQYSRARVFTKAVAV 295
DB	244 QLLHKBELSHVCKWKAFTDKNAPCLDRDRIVECYFWGLSGYEPQYSRARVFTKAVAV 303
QY	296 ITLIDDTYDAGYVEELKIETEAVERNSICTDLTPYMKPIYKLFMDTYTEMEELAKE 355
DB	304 ASIVDDTYDAGYVEELKIETEAVERNSICTDLTPYMKPIYKLFMDTYTEMEELAKE 363
QY	356 GRITDLFCNGKEFYKFNRLNLMVAKWANEHGIPTTEEDHPVVIITGGANLITTCYLGMS 415
DB	364 GROVRVEYAKAMIRLVQVLLKAKWTHQYKPTFEFRNALPTSGYMLAITAFVGMG 423
QY	416 DIETKESVEVAVSAPPLFRYSGLIGRLNLDLMTKAEQERKSSSSLESYMKWYNEEY 475
DB	424 EVITPFTKWAASDPKIKASTICREFNDDIAEHFNHRREDDCSAIECMKYGVTAQE 483
QY	476 AQTLIYKEVEDVWMDINREYLTGNIIPRLMAVILYLCQFLEYQVAGKNFTMGDEYKH 535
DB	484 AYNEFNKHISSMKVDNEEFKPTPEMTPVLCSNLARVMDVLYREGDGYTHVGAAGK 543
QY	536 LIKSLLYVPMIS 547
DB	544 GITSLLDIPQI 555

Search completed: June 3, 2004, 16:41:38
Job time : 47 secs

RESULT 15
Q9FYU6
ID Q9FYU6 PRELIMINARY; PRT; 555 AA.
AC Q9FYU6;
DT 01-MAR-2001 (TRENBLrel. 16, Created)

CC unsuccessful. Sesquiterpenes such as amorphadiene are also useful as
CC flavour and fragrance compounds in the food and perfume industries.
CC Terpenes also play a role in plant-insect interactions, such as the
CC attraction or repulsion of insects by plants. In addition,
CC dihydroartemisinic acid, an intermediate in the metabolic route from
CC amorphadiene to artemisinin in *A. annua*, can be used as an antioxidant.
CC DNA encoding amorphadiene synthase may be used to generate transgenic
CC plants able to synthesise amorphadiene synthase. Preferably, the plants
CC used for amorphadiene production are plants which naturally produce
CC sesquiterpenes, as these plants already have the basic sesquiterpene
CC synthetic pathways and storage compartments. *A. annua* may be transformed
CC with amorphadiene synthase expression constructs, enabling more
CC amorphadiene to be produced and hence increasing artemisin production. It
CC may then be economically feasible to extract artemisin from such plants
CC for use as an antimalarial
CC
XX
SQ Sequence 546 AA;

Query Match 99.5%; Score 2891; DB 3; Length 546;
Best Local Similarity 100.0%; Pred. No. 5.7e-250;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 MALTEKPIRPIANPPSPISWGDQFLIYQKQVEQVQVNDLKKEVRLKALDIPMKH 61
DB 1 MALTEKPIRPIANPPSPISWGDQFLIYQKQVEQVQVNDLKKEVRLKALDIPMKH 60
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
DB 61 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLELLEYATSVMPGVEIILEDALGFTSRSLIMTKDAFS 181
DB 121 VFNNYKDKNGAFKQSLANDVEGLELLEYATSVMPGVEIILEDALGFTSRSLIMTKDAFS 180
QY 182 TNPALFETIORALKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLILKAKLFNLLQSLHKE 241
DB 181 TNPALFETIORALKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLILKAKLFNLLQSLHKE 240
QY 242 ELSHVCNKKWKAQDKNAPCLDRIVECYEWGLSGGYEPOYSRARVFTTRAVAVITLIDD 301
DB 241 ELSHVCNKKWKAQDKNAPCLDRIVECYEWGLSGGYEPOYSRARVFTTRAVAVITLIDD 300
QY 302 TYDAYGYTELKIFTEAVERWSITCLDTPLEVMKPIYKLFMDTYTEMEEFLLAKEGRDLDL 361
DB 301 TYDAYGYTELKIFTEAVERWSITCLDTPLEVMKPIYKLFMDTYTEMEEFLLAKEGRDLDL 360
QY 362 NCGKEFVKZFVRNLMYEAKEWANEHGIPTTEHDHPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB 361 NCGKEFVKZFVRNLMYEAKEWANEHGIPTTEHDHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SYEWAASAPLPFRYSIGLGRRLNDLMTKAEQERKSSSLESVMKEYNVNEEVAQTLIY 481
DB 421 SYEWAASAPLPFRYSIGLGRRLNDLMTKAEQERKSSSLESVMKEYNVNEEVAQTLIY 480
QY 482 KEVEDVWKDINREYLTTKNIPRLLMAVIYLCQFLEVQYAGKDNFTFMGDEYKHLIKSL 541
DB 481 KEVEDVWKDINREYLTTKNIPRLLMAVIYLCQFLEVQYAGKDNFTFMGDEYKHLIKSL 540
QY 542 VYPMIS 547
DB 541 VYPMIS 546

RESULT 2
AAY81272
ID AAY81272 standard; protein; 546 AA.
XX
XX AAY81272;
XX
DT 19-JUN-2000 (first entry)
XX
DE Artemisia annua amorphadiene synthase, encoded by DNA clone 292763.
XX
XX

KW Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;
KW farnearyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;
KW sesquiterpene; flavouring; fragrance.
XX
OS Artemisia annua.
XX EP982404-A1.
XX
PD 01-MAR-2000.
XX
XX 27-AUG-1998; 98EP-00202854.
XX
XX 27-AUG-1998; 98EP-00202854.
XX
PA (WALL/) WALLAART T E.
XX
XX Wallaart TED, Bouwmeester HJ;
XX
DR WPI; 2000-258617/23.
DR N-PSDB; AAZ92763.
XX
XX New isolated DNA sequences and polypeptides comprising amorpha-4,11-diene
PT synthase activity, useful for production of amorphadiene and/or
PT artemisinin.
XX
PS Disclosure; Page 14-16; 41pp; English.

CC The invention relates to DNA encoding Artemisia annua amorpha-4,11-diene
CC synthase (AAZ92754). This enzyme catalyses the cyclisation of farnearyl
CC pyrophosphate (FPP), a ubiquitous precursor of primary and secondary
CC metabolites, into amorpha-4,11-diene. Amorphadiene is a precursor of
CC artemisin, a sesquiterpene lactone endoperoxide produced by *A. annua*
CC which is a promising antimalarial drug candidate. Amorphadiene synthase
CC is a branch point enzyme in the artemisin biosynthetic pathway, and the
CC cyclisation of FPP is therefore likely to be the rate-limiting step.
CC Artemisin is present in very low concentrations in *A. annua*, making it
CC expensive for use as an antimalarial drug, and attempts at organic
CC synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are
CC also useful as flavour and fragrance compounds in the food and perfume
CC industries. Terpenes also play a role in plant-insect interactions, such
CC as the attraction or repulsion of insects by plants. In addition,
CC dihydroartemisinic acid, an intermediate in the metabolic route from
CC amorphadiene to artemisinin in *A. annua*, can be used as an antioxidant.
CC DNA encoding amorphadiene synthase may be used to generate transgenic
CC plants able to synthesise amorphadiene synthase. Preferably, the plants
CC used for amorphadiene production are plants which naturally produce
CC sesquiterpenes, as these plants already have the basic sesquiterpene
CC synthetic pathways and storage compartments. *A. annua* may be transformed
CC with amorphadiene synthase expression constructs, enabling more
CC amorphadiene to be produced and hence increasing artemisin production. It
CC may then be economically feasible to extract artemisin from such plants
CC for use as an antimalarial. The present sequence represents *A. annua*
CC amorphadiene synthase encoded by DNA clone AAZ92763; neither of these
CC sequences are referred to further in the specification

Sequence 546 AA;

Query Match 99.3%; Score 2887; DB 3; Length 546;
Best Local Similarity 99.6%; Pred. No. 1.3e-249;
Matches 544; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MALTEKPIRPIANPPSPISWGDQFLIYQKQVEQVQVNDLKKEVRLKALDIPMKH 61
DB 1 MSLTEKPIRPIANPPSPISWGDQFLIYQKQVEQVQVNDLKKEVRLKALDIPMKH 60
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
DB 61 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLELLEYATSVMPGVEIILEDALGFTSRSLIMTKDAFS 181
DB 121 VFNNYKDKNGAFKQSLANDVEGLELLEYATSVMPGVEIILEDALGFTSRSLIMTKDAFS 180

QY 182 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQDSSHNTLLKLAKLEFNLLQSLHKE 241
DB 181 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQDSSHNTLLKLAKLEFNLLQSLHKE 240
QY 242 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 301
DB 241 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGYEELKIPTFAVERWSITCLDITLPEYMKPIYKLFMDTYTEMEBEFLAKEGRTDLF 361
DB 301 TYDAYGYEELKIPTFAVERWSITCLDITLPEYMKPIYKLFMDTYTEMEBEFLAKEGRTDLF 360
QY 362 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB 361 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVEWAVSAPPLFRYSGLIGRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 481
DB 421 SVEWAVSAPPLFRYSGLIGRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 480
QY 482 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNTRMGDEYKHLIKSLL 541
DB 481 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNTRMGDEYKHLIKSLL 540
QY 542 VYPMSI 547
DB 541 VYPMSI 546

RESULT 3
ID AAM50093 standard; protein; 546 AA.
AC AAM50093;
DT 20-AUG-2002 (first entry)
XX Armopha-4,11-diene synthase associated protein #1.
DE Armopha-4,11-diene synthase; armopha-4,11-diene; antimalaria agent;
KW armosinin biosynthesis.
OS Artemisia annua.
PN KR2001084864-A.
PD 06-SEP-2001.
PF 29-FEB-2000; 2000KR-00010207.
PR 29-FEB-2000; 2000KR-00010207.
PS (KIMS/) KIM S U.
PI Jang YJ, Kim SU, Park SH, Song SH;
XX WPI; 2002-153413/20.
XX
PT Armopha-4,11-diene synthase of Artemisia annua, useful for mediating
PT synthesis of antimalaria agent armopha-4,11-diene.
XX
PS Disclosure; Fig 7; 17pp; Korean.

CC This invention describes a novel armopha-4,11-diene synthase of Artemisia
CC annua, which mediates the synthesis of armopha-4,11-diene in the
CC biosynthesis pathway of artemisinin and is useful as an antimalaria
CC agent. This sequence represents a protein associated with the invention
XX
SQ Sequence 546 AA;
Query Match 99.0%; Score 2878; DB 5; Length 546;
Best Local Similarity 99.3%; Pred. No. 8.4e-249;
Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 MALTEBKPIRPIANFPPSINGDOFLIYKQVQGVQIIVNDLKKEVROLLKEALDIPMKH 61
DB 1 MSUTEKPIRPIANFPPSINGDOFLIYKQVQGVQIIVNDLKKEVROLLKEALDIPMKH 60
QY 62 ANLLKLIIDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTC 121
DB 61 ANLLKLIIDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTC 120
QY 122 VFNNYKDKGAFKQSLANDVEGLELLEYEATSMRVPGBIILBEDALGFTSRLSIMTKDAPS 181
DB 121 VFNNYKDKGAFKQSLANDVEGLELLEYEATSMRVPGBIILBEDALGFTSRLSIMTKDAPS 180
QY 182 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQDSSHNTLLKLAKLEFNLLQSLHKE 241
DB 181 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQDSSHNTLLKLAKLEFNLLQSLHKE 240
QY 242 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 301
DB 241 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGYEELKIPTFAVERWSITCLDITLPEYMKPIYKLFMDTYTEMEBEFLAKEGRTDLF 361
DB 301 TYDAYGYEELKIPTFAVERWSITCLDITLPEYMKPIYKLFMDTYTEMEBEFLAKEGRTDLF 360
QY 362 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB 361 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVEWAVSAPPLFRYSGLIGRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 481
DB 421 SVEWAVSAPPLFRYSGLIGRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 480
QY 482 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNTRMGDEYKHLIKSLL 541
DB 481 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNTRMGDEYKHLIKSLL 540
QY 542 VYPMSI 547
DB 541 VYPMSI 546

RESULT 4
AAY81271
ID AAY81271 standard; protein; 704 AA.
XX AAY81271;
XX 19-JUN-2000 (first entry)
DT
DE Artemisia annua amorphadiene synthase, encoded by a positive DNA clone.
XX Amopha-4,11-diene synthase; amorphadiene synthase; PPP;
KW farnesyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;
KW sesquiterpene; flavouring; fragrance.
XX Artemisia annua.
XX
FH Key Location/Qualifiers
FT Misc-difference 267 /note= "Encoded by CCT"
FT Misc-difference 270 /note= "Encoded by AGA"
FT Misc-difference 272 /note= "Encoded by AGA"
FT Misc-difference 291 /note= "Encoded by AGA"
FT Misc-difference 555 /label= Xaa
FT /note= "Encoded by TGA, Xaa = unknown"
FT Misc-difference 562 /label= Xaa
FT /note= "Encoded by TAG, Xaa = unknown"


```
Misc-difference 569 /note= "Encoded by AGG"  
Misc-difference 596 /label= Xaa  
/note= "Encoded by TAA, Xaa = unknown"  
Misc-difference 613 /label= Xaa  
/note= "Encoded by TAA, Xaa = unknown"  
Misc-difference 640 /label= Xaa  
/note= "Encoded by TGA, Xaa = unknown"  
Misc-difference 651 /label= Xaa  
/note= "Encoded by TGA, Xaa = unknown"  
Misc-difference 696 /label= Xaa  
/note= "Encoded by TAA, Xaa = unknown"
```

EP982404-A1.

01-MAR-2000.

27-AUG-1998; 98EP-00202854.

27-AUG-1998; 98EP-00202854.

(WALL/) WALLAART T E.

Wallaart TED, Bouwmeester HJ:

WPI; 2000-258617/23.

N-PSDB; AAZ92760.

New isolated DNA sequences and polypeptides comprising amorpho-4,11-diene synthase activity, useful for production of amorphadiene and/or artemisinin.

Disclosure; Fig 8A; 41pp; English.

The invention relates to DNA encoding *Artemisia annua* amorphadiene synthase (AAZ92754). This enzyme catalyses the cyclisation of farnesyl pyrophosphate (FPP), a ubiquitous precursor of primary and secondary metabolites, into amorphadiene. Amorphadiene is a precursor of artemisinin, a sesquiterpene lactone endoperoxide produced by *A. annua* which is a promising antimalarial drug candidate. Amorphadiene synthase is a branch point enzyme in the artemisinin biosynthetic pathway, and the cyclisation of FPP is therefore likely to be the rate-limiting step.

Artemisinin is present in very low concentrations in *A. annua*, making it expensive for use as an antimalarial drug, and attempts at organic synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are also useful as flavour and fragrance compounds in the food and perfume industries. Terpenes also play a role in plant-insect interactions, such as the attraction or repulsion of insects by plants. In addition, dihydroartemetic acid, an intermediate in the metabolic route from amorphadiene to artemisinin in *A. annua*, can be used as an antioxidant. DNA encoding amorphadiene synthase may be used to generate transgenic plants able to synthesise amorphadiene synthase. Preferably, the plants used for amorphadiene production are plants which naturally produce sesquiterpenes, as these plants already have the basic sesquiterpene synthetic pathways and storage compartments. *A. annua* may be transformed with amorphadiene synthase expression constructs, enabling more amorphadiene to be produced and hence increasing artemisinin production. It may then be economically feasible to extract artemisinin from such plants for use as an antimalarial. The present sequence represents amorphadiene synthase encoded by a positive cDNA clone which was identified in an *A. annua* cDNA library using a 538 bp amorphadiene synthase probe (AAZ92757) in an exemplification of the present invention.

Sequence 704 AA;

very Match 98.2%; Score 2855; DB 3; Length 704;

1st Local Similarity 99.1%; Pred. No. 1.4e-246;

Matches 541; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	2	MALTEEXPIRDIANPPSISWGQFLIYQKVEQGVQEIIVNDLKCEVRQLLKCALDIPMKH	61
Db	9	MSLTEEXPIRDIANPPSISWGQFLIYQKVEQGVQEIIVNDLKCEVRQLLKCALDIPMKH	68
Qy	62	ANLKLIDEIORLGIPIYHFEREIDHALOCIIYETYGDNWNGRSSLWFLMRKQGYVTCO	121
Db	69	ANLKLIDEIORLGIPIYHFEREIDHALOCIIYETYGDNWNGRSSLWFLMRKQGYVTCO	128
Qy	122	VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPEGIILEDALGFTSRSLSIMTKDAPS	181
Db	129	VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPEGIILEDALGFTSRSLSIMTKDAPS	188
Qy	182	TNPALFTEIQBALQOPLWKRLPRIETAAQIIPFYQQODSHNKTLKLAKLEFNLLQSLHKE	241
Db	189	TNPALFTEIQBALQOPLWKRLPRIETAAQIIPFYQQODSHNKTLKLAKLEFNLLQSLHKE	248
Qy	242	ELSHVCWKWKAFDIKKQAPCLDRDIVECYFWGLSGGYBPQYSRAVEFTTKAVAVITLDD	301
Db	249	ELSHVCWKWKAFDIKKQAFCLPDPPIVECYFWGLSGGYBPQYSRAVEFTTKAVAVITLDD	308
Qy	302	TYDAYGTYEELKIPTFAVERWSITCLDTPLEPMKPIYKLFMDTYTEMEBEFLAKEGRTDLP	361
Db	309	TYDAYGTYEELKIPTFAVERWSITCLDTPLEPMKPIYKLFMDTYTEMEBEFLAKEGRTDLP	368
Qy	362	NCGKEFYKEFVRNLMVEAKWANEHGIPTTEBHDPPVITGGANLLTTTCYLGMSDIFTKE	421
Db	369	NCGKEFYKEFVRNLMVEAKWANEHGIPTTEBHDPPVITGGANLLTTTCYLGMSDIFTKE	428
Qy	422	SVEWAVSAPPLFRYSGILGRNLDMTHKACQERKHSSSSLESYMKENVNNEVEYAQTLIY	481
Db	429	SVEWAVSAPPLFRYSGILGRNLDMTHKACQERKHSSSSLESYMKENVNNEVEYAQTLIY	488
Qy	482	KEVEDWVKDINREYLTTKNIPRPLLMVAYILYCQFLEVQYAGKDNFTRMGDEYKHLIKSLL	541
Db	489	KEVEDWVKDINREYLTTKNIPRPLLMVAYILYCQFLEVQYAGKDNFTRMGDEYKHLIKSLL	548
Qy	542	VYPMISI 547	
Db	549	VYPMISI 554	

RESULT 5

AAM50094

ID AAM50094 standard; protein; 547 AA.

AC AAM50094;

DT 20-AUG-2002 (first entry)

DE Armopha-4,11-diene synthase associated protein #2.

Armopha-4,11-diene synthase; armopha-4,11-diene; antimalaria agent;
KW
artemisinin biosynthesis.
KW

OS Artemisia annua.

PN KR2001084864-A.

PD 06-SEP-2001.

29-FEB-2000; 2000KR-00010207.

PR 29-FEB-2000; 2000KR-00010207.

PA (KIMS/) KIM S U

PI Jang YJ, Kim SU, Park SH, Song SH;

DR WPI; 2002-153413/20.

Armopha-4, 11-diene synthase of *artemisia annua*, useful for mediating
PT synthesis of antimalaria agent armopha-4, 11-diene.
PT

Query Match	38.6%;	Score 1123;	DB 2;	Length 548;
Best Local Similarity	38.8%;	Pred. No. 1.9e-91;		
Matches 214;	Conservative 123;	Mismatches 206;	Indels 8;	Gaps 3;
QY	2	MALTEEKPIRPIANPPSPISGQFLIYQKQVGEQVINDLKEVRQLLKALDIPMKH	61	
DB	1	MAASADKCRPLANFHPVSWGHHFSLYTHETINQEKVEYDEYKIRKMLVETCD---NS	57	
QY	62	ANLLKLIDEIQRGLPIYHFEREIDHALQCIYETYGDNNWNGDRS----SLWFLMRKQGY	117	
DB	58	TQKLVLDAMQRLGVAYHFDNEIETSIQNIIFDASSKQNDNNLVVSLRFLVRQGGHY	117	
QY	118	VTCDVFNKYKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRLSIMTK	177	
DB	118	MSSDVFQKFTNQDGKFKETLTNDVQGLLSLYEASHLRVNEEILBEALFTTTHLESIVS	177	
QY	178	DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQDSHNKTLKLAKLEFNLOS	237	
DB	178	NLSNNNSLKVEVGALTPQIRMTLPRMGARKYISYIENNDAAHLLLFKALDFNMLQK	237	
QY	238	LHKESLHVCKWKAFFDKKAPCLRDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT	297	
DB	238	PHQRELSDLTRWKKDLDFANKYPYARDRLVECYFWILGVYFEPKYSRARKMTKVLNLT	297	
QY	298	LIDDTYDAYGTVEELKIETEAVERHSITCLDPLPYMKPIYKLFMDTYTEMEFLAKEGR	357	
DB	298	IIDTDFAYATFDELVTNDAIQRWDANAIDSIOQYMRPAYQALLDIYSEMQVLSKEGK	357	
QY	358	TDLFCNGKEFVKFEFVRNLVMEAKWANE-GHPTTEHDPVVIITGGANLLTTCYLGMSD	416	
DB	358	LDRVYAKNEMKLVRAVFKETQWLNDCHDHPKYEQVENAIVSAGYMMISTTCLVGIEE	417	
QY	417	IFTKESVEMAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSSSSLESYMKYNNVEYA	476	
DB	418	FISHTFELWNNESVIVRASALIRAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET	477	
QY	477	QTLIYKEVDYKNDINREYLTNTNPRLLMAVILCQLEFVQYAGKDNFTMGDYKHL	536	
DB	478	YIKELKEVTNAWKNDINKQFRTPEVFMFLYERVLNLTRVADTLYREKDTYSTAKGLKNN	537	
QY	537	IKSLLVYPMIS	547	
DB	538	INPLIESVKI	548	
RESULT 9				
ID	AA90856	standard; protein; 548 AA.		
AC	AA90856;			
XX				
DT	25-AUG-2000	(first entry)		
XX				
DE				
XX				
KW				
KW				
KW				
KW				
KW				
KW				
OS				
XX				
PN				
XX				
PD				
PF				
XX				
PR				
PR				

PR	23-AUG-1999;	99US-0150262P.		
XX				
PA	(KENT) UNIV KENTUCKY RES DEPT.			
PA	(SALK) SALK INST BIOLOGICAL STUDIES.			
XX				
PI	Chappell J, Manna KR, Noel JP, Starks CM;			
XX				
DR	WPI; 2000-232839/25.			
DR	N-PSDB; AAA38935.			
XX				
PT	Novel terpene synthase enzymes, useful for producing terpene			
PT	hydrocarbons, e.g. fragrances or antitumor agents, are derived from known			
PT	enzymes by specific amino acid alterations.			
XX				
XX	Claim 111; Page 432-434; 450pp; English.			
XX				
CC	The present invention describes an isolated terpene synthase (I)			
CC	comprising a region with at least 20% identity to region 265-535 of a 548			
CC	amino acid (aa) sequence (Ia), given in AA90831. (I) contains nine alpha			
CC	-carbon atoms (alphaC) that have interatomic distances, between each			
CC	other, within tabulated ranges, have a centre point (within a sphere of			
CC	radius 2.3 Angstrom) within tabulated ranges, and have an ordered			
CC	arrangement of R groups (defining aa side chains), excluding specific			
CC	tabulated arrangements (tables given in the specification). (I), and			
CC	related enzymes, are used to produce a wide range of terpenoids (e.g.			
CC	cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,			
CC	flavours, pheromones, defensive agents, pigments, antitumor agents,			
CC	components of signal transduction pathways, precursors of steroid			
CC	hormones and bile acids, as photoreceptors and as co-factor side chains.			
CC	Some synthases with little or no catalytic activity (and nucleic acids			
CC	encoding them) are used as controls in the analysis of products formed by			
CC	enzymatic synthesis; as nutrient supplements; for affinity purification			
CC	of isoprenoids; or to develop immunological reagents or nucleic acids for			
CC	monitoring expression of terpene synthase or inheritance of the gene in			
CC	plant breeding programs. The new synthases may produce novel terpene			
CC	products. AAA38910 to AAA38938 and AA90831 to AA90859 represent			
CC	sequences used in the exemplification of the present invention			
XX				
SQ	Sequence 548 AA;			
	Query Match	38.6%;	Score 1123;	DB 3; Length 548;
	Best Local Similarity	38.8%;	Pred. No. 1.9e-91;	
	Matches 214;	Conservative 123;	Mismatches 206;	Indels 8; Gaps 3;
QY	2	MALTEEKPIRPIANPPSPISGQFLIYQKQVGEQVINDLKEVRQLLKALDIPMKH	61	
DB	1	MAASADKCRPLANFHPVSWGHHFSLYTHETINQEKVEYDEYKIRKMLVETCD---NS	57	
QY	62	ANLLKLIDEIQRGLPIYHFEREIDHALQCIYETYGDNNWNGDRS----SLWFLMRKQGY	117	
DB	58	TQKLVLDAMQRLGVAYHFDNEIETSIQNIIFDASSKQNDNNLVVSLRFLVRQGGHY	117	
QY	118	VTCDVFNKYKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRLSIMTK	177	
DB	118	MSSDVFQKFTNQDGKFKETLTNDVQGLLSLYEASHLRVNEEILBEALFTTTHLESIVS	177	
QY	178	DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQDSHNKTLKLAKLEFNLOS	237	
DB	178	NLSNNNSLKVEVGALTPQIRMTLPRMGARKYISYIENNDAAHLLLFKALDFNMLQK	237	
QY	238	LHKESLHVCKWKAFFDKKAPCLRDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT	297	
DB	238	PHQRELSDLTRWKKDLDFANKYPYARDRLVECYFWILGVYFEPKYSRARKMTKVLNLT	297	
QY	298	LIDDTYDAYGTVEELKIETEAVERHSITCLDPLPYMKPIYKLFMDTYTEMEFLAKEGR	357	
DB	298	IIDTDFAYATFDELVTNDAIQRWDANAIDSIOQYMRPAYQALLDIYSEMQVLSKEGK	357	
QY	358	TDLFCNGKEFVKFEFVRNLVMEAKWANE-GHPTTEHDPVVIITGGANLLTTCYLGMSD	416	
DB	358	LDRVYAKNEMKLVRAVFKETQWLNDCHDHPKYEQVENAIVSAGYMMISTTCLVGIEE	417	
QY	417	IFTKESVEMAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSSSSLESYMKYNNVEYA	476	

Db 418 FISHTEFFELMESVIVRASALIARANDIVGHEDEQERGHVASLIECYMKDYGASKQET 477
 Qy 477 QFLIYKEVEDVKDINREYLTTRNIPRLPMVILYCOFLEVYAGKDNFRMGDEYKHL 536
 Db 478 YIKFLKEVTNAWKDINKQFSPREVEFMFVLNLRVADILYKEDTYSYTAGKLNKM 537
 Qy 537 IKSLIVYPMXI 547
 Db 538 INPILIESVKI 548

RESULT 10

AD82708
 ID ADE82708 standard; protein; 556 AA.
 AC ADE82708;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Terpenoid biosynthesis related H64 strawberry protein #14.
 XX
 KW isoprenoid; bio-active compound synthesis; pesticide; dermatological;
 KW cytostatic; immunosuppressive; virucide; flavour; fragrance;
 KW bio-control agent; food additive; food industry; pest control;
 KW degreasing solvent; plasticizer; dye carrier; dental caries;
 KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
 KW anti-retroviral; monoterpene alcohol linalool;
 KW sesquiterpene alcohol nerolidol; monoterpene; strawberry.
 XX
 OS Fragaria x ananassa.
 XX
 PN EP1231273-A1.
 XX
 PD 14-AUG-2002.
 XX
 PF 12-FEB-2001; 2001EP-00200488.
 XX
 PR 12-FEB-2001; 2001EP-00200488.
 XX
 PA (PLAN-) PLANT RES INT BV.
 XX
 PI Aharoni A, Verhoeven HA, Jongema MA, Bouwmeester HJ;
 XX
 DR WPI; 2003-879727/82.

Novel recombinant nucleic acid encoding proteinaceous molecule, useful for producing flavor, fragrance and/or biocontrol agent which is useful as food additive in processed food industry and as antimicrobial agent.

Disclosure; Page; 52pp; English.

The invention relates to a novel isolated or recombinant nucleic acid or its functional fragment, encoding a proteinaceous molecule essentially capable of isoprenoid bio-active compound synthesis when provided with a suitable substrate under appropriate reaction conditions. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid and its protein have the following activities: pesticide, dermatological, cytostatic, immunosuppressive, and virucide. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid is useful for producing flavour, fragrance, and/or a bio-control agent, by transforming or transfecting a suitable host with the recombinant isoprenoid bio-active compound synthesis nucleic acid, expressing the recombinant isoprenoid bio-active compound synthesis nucleic acid in the presence of a suitable substrate, and optionally isolating the formed product. The bio-control agent is useful as an anti-microbial agent, as a food additive in the processed food industry to modify the taste of syrups, ice-creams, frozen desserts, yogurts, confectionary and like products, as a flavouring agent for oral medications and vitamins, and for providing additional flavour/aroma in beverages, including alcoholic beverages. The bio-control agent is also useful for enhancing or reducing flavour, aroma, fragrance or scent of plants, natural products, and/or synthetic or artificial products, and for the industrial synthesis of nature

CC identical flavour/aroma substances and/or artificial flavour/aroma
 CC substances. The bio-control agent is also useful as a pest control agent
 CC for the biological control of the interaction between plants and insects
 CC and/or plants and microorganisms, for providing flavour/aroma in
 CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
 CC products, personal care products and health care products, as a
 CC disinfectant additive and in the preparation of a composition. The novel
 CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
 CC fragments is useful as a molecular marker or diagnostic tool. The protein
 CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
 CC acid is useful for the production of an antagonist e.g. an antibody or
 CC its functional equivalent which is useful for inhibiting the synthesis of
 CC the bio-control agent. A composition, containing the bio-control agent,
 CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
 CC the aroma and/or taste of food or non-food products, and/or protection of
 CC food or non-food products against fungal contamination and/or pest
 CC infestation. The composition is also useful for the biological control of
 CC pests, for the protection of stored products and for the prevention or
 CC treatment of disease. The bio-control agent is useful as a degreasing
 CC solvent, plasticizer and dye carrier. The composition is useful for
 CC replacing potentially carcinogenic synthetic food additives currently
 CC used. The composition is also useful for treating dental caries, dental
 CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and
 CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
 CC compound synthesis nucleic acid or its protein is useful for the
 CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
 CC nerolidol, and monoterpene. This sequence represents an H64 protein
 CC used in the terpenoid biosynthesis method of the invention. NOTE: This
 CC sequence is not shown in the specification. It has been obtained from
 CC electronic data supplied with this specification from the European Patent
 CC Office.

XX Sequence 556 AA;

Query Match 38.4%; Score 1115.5; DB 7; Length 556;
 Best Local Similarity 40.9%; Pred. NO. 9.1e-91;
 Matches 226; Conservative 113; Mismatches 195; Indels 19; Gaps 5;

Qy 4 LEEKPIRPIANFPSPISWGDQF-----LIYQKQVEQVEQIVNLKKEVRQLKALDI 57
 Db 14 ISKPEVVRTANFPSPISWGDQF-----LIYQKQVEQVEQIVNLKKEVRQLKALDI 57
 Qy 58 PMKHANLLKLIDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGD---RSSLWFLMRKQ 114
 Db 70 --DSHQKLKLIDEIQRLGVAYHFESEIDQALRIHETQDIHDGGDLNVALRFRLLRRH 127
 Qy 115 GYVTCDFVNNYKDKNGAFKOSLANDVEGLLELYEATSNRVPEGEILEDALGFTRRRLSI 174
 Db 128 GYNVSCDFVNFKDTNGDYKKSLSVTDLSGMLSFYEAHLRVHGEKLLLEALVFTTHL-- 185
 Qy 175 MTKDAFTNPALFTEIQALQKPLKRLPRIEAAQYIPYQQQDSHNKTLKLAKLEFNL 234
 Db 186 --QSASAKSSLLKQTITAEVERPLKTKWELGARRYSIYQDEASYNLLKLAKLDFNV 243
 Qy 235 LQSLHKELSHVCKWKAFAIDIKNAPCLDRIVECYFWGLSGSYEPQYSRARVFTTKAVA 294
 Db 244 VQCLHKELSILLRWYKELDFARRPFADRIVELFFWLAGIYFEPYVFGRIHLAKLIE 303
 Qy 295 VITLIDDDYDAGTYEELKIFTEAVERMSITCLDTLPYMKPIYKLFMTYTMESFLAK 354
 Db 304 ITTVMDNDYDAFGTPEELVILTEALDRWDASCMDQLPDMQPFYITLDDVIDEVEBELTK 363
 Qy 355 EGRDLPNCGKGFVKFVFNLMVKAWEKNEGHIPTTEHDPVVIITGGANLLTTCVLGM 414
 Db 364 QGRSVRIHYAKEMKQARLYFAEAIFWHEGCTPKMDGVMRVAASVGNMTLSVSLVGM 423
 Qy 415 SDIFTKESVWAVSAPPLFRYSGLILGRRLNDLMTKAEQKHSSSLSLEYMKYNNVEE 474
 Db 424 GDILTKPEFELWLTNPKILRASNTIFRLMDDDIAGYKFEKRGHVASSIDCYWNEYGSEQ 483
 Qy 475 YAOQLIYKEVEDVKDINREYLTTRNIPRLPMVILYCOFLEVYAGKDNFRMGDEYK 534
 Db 484 ETIDIFNKRIVDWSDKIDNEFLRPTAAFPVFLNRLNLRVVDLLYKRGDAFTHVGLMK 543

QY 535 HLIKSLVYPMSI 547
Db 544 DCIAAMPIDPVL 556

RESULT 11
ADE82706
ID ADE82706 standard; protein; 555 AA.
XX AC ADE82706;
XX DT 29-JAN-2004 (first entry)
XX DE Terpenoid biosynthesis related H64 strawberry protein #13.
XX KW isoprenoid; bio-active compound synthesis; pesticide; dermatological;
KW cytosatic; immunosuppressive; virucide; flavour; fragrance;
KW bio-control agent; food additive; food industry; pest control;
KW degreasing solvent; plasticizer; dye carrier; dental caries;
KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
KW anti-retroviral; monoterpene alcohol linalool;
KW sesquiterpene alcohol nerolidol; monoterpene; strawberry.
XX OS Fragaria x ananassa.
XX PN EP1231273-A1.
XX PD 14-AUG-2002.
XX PF 12-FEB-2001; 2001EP-00200488.
XX PR 12-FEB-2001; 2001EP-00200488.
XX PA (PLAN-) PLANT RES INT BV.
XX PI Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;
XX WP; 2003-879727/82.
XX PT Novel recombinant nucleic acid encoding proteinaceous molecule, useful
PT for producing flavor, fragrance and/or biocontrol agent which is useful
PT as food additive in processed food industry and as antimicrobial agent.
PS Disclosure; Page: 52pp; English.

The invention relates to a novel isolated or recombinant nucleic acid or its functional fragment, encoding a proteinaceous molecule essentially capable of isoprenoid bio-active compound synthesis when provided with a suitable substrate under appropriate reaction conditions. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid and its protein have the following activities: pesticide, dermatological, cytosatic, immunosuppressive, and virucide. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid is useful for producing flavour, fragrance, and/or a bio-control agent, by transforming or transfecting a suitable host with the recombinant isoprenoid bio-active compound synthesis nucleic acid, expressing the recombinant isoprenoid bio-active compound synthesis nucleic acid in the presence of a suitable substrate, and optionally isolating the formed product. The bio-control agent is useful as an anti-microbial agent, as a food additive in the processed food industry to modify the taste of syrups, ice-creams, frozen desserts, yogurts, confectionary and like products, as a flavouring agent for oral medications and vitamins, and for providing additional flavour/aroma in beverages, including alcoholic beverages. The bio-control agent is also useful for enhancing or reducing flavour, aroma, fragrance or scent of plants, natural products, and/or synthetic or artificial products, and for the industrial synthesis of nature identical flavour/aroma substances and/or artificial flavour/aroma substances. The bio-control agent is also useful as a pest control agent for the biological control of the interaction between plants and insects and/or plants and microorganisms, for providing flavour/aroma in cosmetics, creams, sun-protectant products, hair conditioners, cleaning products, personal care products and health care products, as a

CC cinntractant additive and in the preparation of a composition. The novel CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its CC fragments is useful as a molecular marker or diagnostic tool. The protei CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic CC acid is useful for the production of an antagonist e.g. an antibody or CC its functional equivalent which is useful for inhibiting the synthesis of CC the bio-control agent. A composition, containing the bio-control agent, CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing CC the aroma and/or taste of food or non-food products, and/or protection of CC food or non-food products against fungal contamination and/or pest CC infestation. The composition is also useful for the biological control of CC pests, for the protection of stored products and for the prevention or CC treatment of disease. The bio-control agent is useful as a degreasing CC solvent, plasticizer and dye carrier. The composition is useful for CC replacing potentially carcinogenic synthetic food additives currently CC used. The composition is also useful for treating dental caries, dental CC plaque and skin disorders, and for immunosuppressive, anti-leukemia and CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active CC compound synthesis nucleic acid or its protein is useful for the CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol CC nerolidol, and monoterpene. This sequence represents an H64 protein CC used in the terpenoid biosynthesis method of the invention. NOTE: This CC sequence is not shown in the specification. It has been obtained from CC electronic data supplied with this specification from the European Patent CC Office.
XX SQ Sequence 555 AA;
Query Match 37.9%; Score 1100; DB 7; Length 555;
Best Local Similarity 41.1%; Pred. No. 2.2e-89;
Matches 225; Conservative 110; Mismatches 192; Indels 20; Gaps 6

Qy 10 IRPIANPPSPISMGDQP-----LIVQXVEGVGQIVNDLKKEVRQLKEALDIPMKUAN 63
Db 20 VRTANFKPSVGDRFANYAEDIIITQMQRQVBELKQVRKEV---FTNAAD----DSSH 73
Qy 64 LKLIDEIORLGIPYHFERBIDHALQCITYEYGNMNGD---RSLWFLRMKGQYVYC 120
Db 74 QLKUIDEIORLGVAIHFESEIDQALERIHETIQDIHGDDYNNVALRFLRRHGYNVSC 133
Qy 121 DVFNKYDKQGAFKQS LANDVEGLLELYEATSMRVPGSEIILEDALGETRSRLSINTKD AF 180
Db 134 DVFNFKDTNGDYKKSLVTDLSGMLSFYEAHLAVHGEKLEALVFTTHL----QSAS 189
Qy 181 STNPALFTEIQRALKQPLWKRLPRIABAAQVIPFQQODSHNKTLCLKAKLFNLLQSLHK 240
Db 190 AKSSLLKTQITEAVER--LLKTMERLGARRYMSIYQDEASYSENILKLAKLDFFNVQCLHK 248
Qy 241 EELSHCVCKWKAPEDKGNAPCLCDRIVECVFGWSGVGPQYSRARVFFTKAVAVITLID 300
Db 249 KELSDILRWYKELDFAARMFPARDRIVELFWTAGIYFEPYVPFRHLTKLIEITVMD 308
Qy 301 DTVDAYGTYTEELKI FTBAVERWSITCIDLTFEYMKPIYKLFMDTYTMEEF LAKEGR TDL 360
Db 309 DMWDAGFTGFEELVNLTAEIRDWDASCMDQLPDYMPQPVITLLDIVDEEEBELTKQGRSYR 368
Qy 361 FNCCKEYKFVRNLVWEAKWANECHPTTEEHDHPVIIIGGANLLTTCTYGLNSDIFPK 420
Db 369 IHYAKETIMQNQARLYFAEARPFHEGCTPRKDEYMRVAASSVGNFMLSVSVLVGMGDITK 428
Qy 421 ESVEWAYSA PPLFRYSOILGRRLNDLMTHKAEORKHSSSLESYMKCEYNNVEYAQT LI 480
Db 429 FEFEWLNEPKILRASNTIFRLMDDIAGYKFEKERGHVASIDCYMNEYGVSBQETIDIF 488
Qy 481 KYEVEDVMKDINREYLTTKNIPRPLLMAVTYLCQFLEVQVYAGKDNFTRMGDEYKHLIKSL 540
Db 489 NKRIVDSWKDINEBLRPATAAPVVLNRVLNTRVVDLLYKRGDAFTHVGLKMLKDCIAAM 548
Qy 541 LVYPMISI 547
Db 549 FIDFVPL 555

DT XX 25-AUG-2000 (first entry)

DE XX Potato vetaspiradiene synthase protein sequence SEQ ID NO:32.

XX Synthese; protein co-ordinate data; active site; modification; terpenoid;

KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;

KW isoprenoid; breeding programme; fragrance; flavour; pheromone;

KW defensive agent; pigment; antitumour; steroid hormone;

KW signal transduction pathway; bile acid; affinity purification;

KW photoreceptor; enzymatic synthesis; nutrient supplement;

XX immunological reagent.

OS Solanum tuberosum.

XX W0200017327-A2.

XX 30-MAR-2000.

XX 17-SEP-1999; 99WO-US021419.

XX 18-SEP-1998; 98US-0100993P.

PR 22-APR-1999; 99US-0130628P.

PR 23-AUG-1999; 99US-0150262P.

XX (KENT) UNIV KENTUCKY RES DEPT.

PA (SALK) SALK INST BIOLOGICAL STUDIES.

XX Chappell J, Manna KR, Noel JP, Starks CM;

XX WPI; 2000-292839/25.

DR N-PSDB; AAA38928.

XX Novel terpene synthase enzymes, useful for producing terpene

PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known

PT enzymes by specific amino acid alterations.

PS Claim 118; Page 395-396; 450pp; English.

XX The present invention describes an isolated terpene synthase (I)

CC comprising a region with at least 20% identity to region 265-535 of a 548

CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha

CC -carbon atoms (alphaC) that have interatomic distances, between each

CC other, within tabulated ranges, have a centre point (within a sphere of

CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered

CC arrangement of R groups (defining aa side chains), excluding specific

CC tabulated arrangements (tables given in the specification). (I), and

CC related enzymes, are used to produce a wide range of terpenoids (e.g.

CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,

CC flavours, pheromones, defensive agents, pigments, antitumour agents,

CC components of signal transduction pathways, precursors of steroid

CC hormones and bile acids, as photoreceptors and as co-factor side chains.

CC Some synthases with little or no catalytic activity (and nucleic acids

CC encoding them) are used as controls in the analysis of products formed by

CC enzymatic synthesis; as nutrient supplements; for affinity purification

CC of isoprenoids; or to develop immunological reagents or nucleic acids for

CC monitoring expression of terpene synthase or inheritance of the gene in

CC plant breeding programs. The new synthases may produce novel terpene

CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent

CC sequences used in the exemplification of the present invention

XX Sequence 556 AA;

SQ Query Match 35.8%; Score 1041; DB 3; Length 556;

Best Local Similarity 39.6%; Pred. No. 4.4e-84;

Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 BEKPIRPIANPPSTWGQDFLIY--QKQVEQGEQVQVNDIKKEVROLKKEALDIPMKHAN 63

DB 13 EEEIVRPADSPSLWGRFHFSDLDNQAGYAOEITLKQSRILLSASSRTU--AE 70

QY 64 LLKLIDEIQRLCIPYHFEIRIDHALQCIYET-----YGDNWNNGDRSLNFRMRKOG 115

DB 71 KLDLIDIVERLUGIAHFKEQKIDMDLDFQYKADPNFEAHEYNDL---QTLVSQFRLLRQHG 127

QY 116 YYTCDVFNXYKDNKNGAFKQSLANDVEGLLELYEATSMRVGEIILEDALGFTSRSLSTM 175

DB 128 YNISPFLKFIHQDAKGFKESLNCNDIKGLNLNLYEASHVTRTHGEDILEESALAFSTAHLE-- 185

QY 176 TKDAFSTNPALFTFIORALKQPLWKRLPRIEAAQVIPPYQOODSHNKTLILKLAKLEFNLL 235

DB 186 -SAPHLKSPSLKQVTHALQSLHKSIPRVETRYFISIEEEOQKNDVLLQPAKLDNFLL 244

QY 236 QSLHKBELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFFTKAVAV 295

DB 245 QMLHKQELSESVSRWKOLDFTVTLPLFARDAVECYFTMTGVYAEPPQYSGARVNLAKTIAM 304

QY 296 ITLIDDTYDAYGNYEELKIETEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMBEFLAKE 355

DB 305 ISIVDDTFDAYGIVKELEIYTDALQWRWDISQIDRLPDYMKLSYKALLDLYNDYEMELSD 364

QY 356 GRDTLFCGKEFVKFVVRNLMVEAKWANEHGIPTTEBHDPPVVIITGGANLLTTTTCYLGM 415

DB 365 GRSDVVHYAKERKEIVRNYFVEAKWFIEGVMPVSEVLSNALATSTYVLLTTTTLGLMK 424

QY 416 DIFTKSSVEWASAPPLFRYSGILGRRLNDLMTKAKQERKHSSSSLSYMKYNNBEY 475

DB 425 SA-NKQDFEWLAKNPKILEANVTLCRVDDIATYVEVKRGQIATGIECYMRDYGVSTEK 483

QY 476 AQTLLYKEVEDVKNKDNREYLTKNIIPRLIMAVIYLCQFLEVOYA-GKDKFTRMGDEYK 534

DB 484 AMEKFOEMASTAWKDVNEGILRPTPVSTETILTRILNLAIRIDVTYKKNQDGYTHPEKVJK 543

QY 535 HLIKSLLVFPMIS 547

DB 544 PHIALLVDSIEI 556

RESULT 14

AAY90831

ID AAY90831 standard; protein; 548 AA.

XX AC AAY90831;

XX DT 25-AUG-2000 (first entry)

DE Tobacco 5-epi-aristolochene synthase protein SEQ ID NO:2.

XX Synthese; protein co-ordinate data; active site; modification; terpenoid;

KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;

KW isoprenoid; breeding programme; fragrance; flavour; pheromone;

KW defensive agent; pigment; antitumour; steroid hormone;

KW signal transduction pathway; bile acid; affinity purification;

KW photoreceptor; enzymatic synthesis; nutrient supplement;

XX immunological reagent.

XX Nicotiana sp.

XX W0200017327-A2.

XX 30-MAR-2000.

XX 17-SEP-1999; 99WO-US021419.

XX 18-SEP-1998; 98US-0100993P.

PR 22-APR-1999; 99US-0130628P.

PR 23-AUG-1999; 99US-0150262P.

XX (KENT) UNIV KENTUCKY RES DEPT.

PA (SALK) SALK INST BIOLOGICAL STUDIES.

XX Chappell J, Manna KR, Noel JP, Starks CM;

XX WPI; 2000-292839/25.

DR N-PSDB; AAA38910.

XX Novel terpene synthase enzymes, useful for producing terpene

१

Search completed: June 3, 2004, 16:40:46
Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 16:30:26 ; Search time 23 Seconds
(without alignments)
1232.289 Million cell updates/sec

Title: US-09-763-822A-14
Perfect score: 2906
Sequence: 1 TVALTEKPIRPIANFPPI.....GDEYKHLIKSLVPMSTGS 549

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1135	39.1	548	US-09-601-091-2	Sequence 2, Appli
2	1123	38.6	548	US-09-601-091-4	Sequence 4, Appli
3	1123	38.6	548	US-09-398-395A-52	Sequence 52, Appl
4	1123	38.6	548	US-09-887-586A-52	Sequence 52, Appl
5	1123	38.6	548	US-09-895-752-52	Sequence 52, Appl
6	1123	38.6	548	US-09-903-012B-52	Sequence 52, Appl
7	1123	38.6	548	US-09-900-797-52	Sequence 52, Appl
8	1041	35.8	556	US-09-398-395A-32	Sequence 32, Appl
9	1041	35.8	556	US-09-887-586A-32	Sequence 32, Appl
10	1041	35.8	556	US-09-895-752-32	Sequence 32, Appl
11	1041	35.8	556	US-09-903-012B-32	Sequence 32, Appl
12	1041	35.8	556	US-09-900-797-32	Sequence 32, Appl
13	1020.5	35.1	548	US-09-398-395A-2	Sequence 2, Appli
14	1020.5	35.1	548	US-09-887-586A-2	Sequence 2, Appli
15	1020.5	35.1	548	US-09-895-752-2	Sequence 2, Appli
16	1020.5	35.1	548	US-09-903-012B-2	Sequence 2, Appli
17	1020.5	35.1	548	US-09-900-797-2	Sequence 2, Appli
18	1017.5	35.0	548	US-09-398-395A-8	Sequence 8, Appli
19	1017.5	35.0	548	US-09-887-586A-8	Sequence 8, Appli
20	1017.5	35.0	548	US-09-895-752-8	Sequence 8, Appli
21	1017.5	35.0	548	US-09-903-012B-8	Sequence 8, Appli
22	1017.5	35.0	548	US-09-900-797-8	Sequence 8, Appli
23	1010.5	34.8	548	US-09-398-395A-6	Sequence 6, Appli
24	1010.5	34.8	548	US-09-887-586A-6	Sequence 6, Appli
25	1010.5	34.8	548	US-09-895-752-6	Sequence 6, Appli
26	1010.5	34.8	548	US-09-903-012B-6	Sequence 6, Appli
27	1010.5	34.8	548	US-09-900-797-6	Sequence 6, Appli

28	1009.5	34.7	548	US-09-398-395A-12	Sequence 12, Appl
29	1009.5	34.7	548	US-09-887-586A-12	Sequence 12, Appl
30	1009.5	34.7	548	US-09-895-752-12	Sequence 12, Appl
31	1009.5	34.7	548	US-09-903-012B-12	Sequence 12, Appl
32	1009.5	34.7	548	US-09-900-797-12	Sequence 12, Appl
33	1005.5	34.6	550	US-08-443-639-8	Sequence 8, Appli
34	1005.5	34.6	550	US-08-577-483-8	Sequence 8, Appli
35	1005.5	34.6	550	US-09-435-380-8	Sequence 8, Appli
36	1001.5	34.5	548	US-09-398-395A-10	Sequence 10, Appl
37	1001.5	34.5	548	US-09-887-586A-10	Sequence 10, Appl
38	1001.5	34.5	548	US-09-895-752-10	Sequence 10, Appl
39	1001.5	34.5	548	US-09-903-012B-10	Sequence 10, Appl
40	1001.5	34.5	548	US-09-900-797-10	Sequence 10, Appl
41	1000.5	34.4	548	US-09-398-395A-4	Sequence 4, Appli
42	1000.5	34.4	548	US-09-887-586A-4	Sequence 4, Appli
43	1000.5	34.4	548	US-09-895-752-4	Sequence 4, Appli
44	1000.5	34.4	548	US-09-903-012B-4	Sequence 4, Appli
45	1000.5	34.4	548	US-09-900-797-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-601-091-2
; Sequence 2, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073, 579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-601-091-2

Query Match 39.1%; Score 1135; DB 4; Length 548;
Best Local Similarity 39.2%; Pred. No. 1.3e-102;
Matches 216; Conservative 122; Mismatches 205; Indels 8; Gaps 3;
QY 2 MALTTEKPIRPIANFPPISTWGOFLYQKQVQGVQIYNDLKKVQQLKKEALDIPMKH 61
Db 1 MAASSADKCRPLANFHPSPVGVYHFLSYTHEITNQEKVEVDEYKETIRKMLVETCD--NS 57
QY 62 ANLLXLIIDEIORLIGIPYHFEREIDHALQCTIYETYGDNWNGDRS-----SLAFRLMKQGY 117
Db 58 TOKVLIDAMQRLGVAYHFDNEIETSIQIFDASSKQNDNDNDYVVSRLFRVQOQHY 117
QY 118 VTCDVFNNYKDKNGAFKQSLANDVEGLELLEATSRVPGCEIILEDALGFTSRSLRSLMTK 177
Db 118 MSSDVFQKQTNQDGRKFKELTNDVQGLLSIYEASHLRVRNEELIEBALTTTTHLESIVS 177
QY 178 DAFSTNPALFTBIQIRALKQPLWKRLPRIAAQVIPPYQOODSHNKTLKLALEFNLLQS 237
Db 178 NLSNNNSLKVGEALTFQIRMTLPRMGARKYISYENNDAAHHLLKFKAKLDFNMLQK 237
QY 238 LHKEELSHVCKWKAFDIKKNAPCLRDRIVECVFVGLSGSYEPQYSRVRVFTTKAVAIT 297
Db 238 FQRELSLUTRWKOLDLFANKYPADRLVECFWILGVYFEPKYSRARKOMTKVLNLT 297
QY 298 LIIDYDAYGTVEELKIFTEAVERWSITCLDTLPEYMKPKLYKLFMDTYTEWEEFLAKEGR 357
Db 298 IIDDFFDAVATFDELVTFNDAIQRDNAIDSIQPYRPAQALLDIYSEWQVLSKEGK 357

QY 477 QTLIYKEVDVWKDINREYLTTKNIPRLLMAVILCOFLEVOYAGKDNFTMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRPTVEPMFVLRLNLRVADTLYKEKDITYSTAKGLKXNM 537
QY 537 IKSLLVYPMISI 547
Db 538 INPILESVKI 548

RESULT 4
US-09-887-586A-52
; Sequence 52, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-887-586A-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;
QY 2 MALTEEKPIRPIANFPSPISWGDQFLIYKQVEQVQVNDLKEVRQLLKEALDIPMKH 61
Db 1 MAASADKCRPLANFHSVWGHFLSYTHETNQKEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETTYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNI FDASSKQNDNDNNLVVSLRFLVRQGHY 117
QY 118 VTCDVFNKYKNGAPKQSLANDVEGLELLELYEATSMRVPGEIILEDALGFTSRSLIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSLYEASHLRVNEEILEEALTFTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEPNLQOS 237
Db 178 NLSNNNSLKVEGEALTQPTMTLPRMGARKYISYIENNDHAHLLLKFAKLDENMLQK 237
QY 238 LHKEELSHVCXKWKAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT 297
Db 238 FHQRELSDLTRWKLDPANKYPYARDRLVECYFWILGVYFEPKYSRARKWMTKVLNLT 297
QY 298 LIDDTYDAYGTYEELKIITEAVERWSITCLDTPYMKPIYKLFMDITYTEMEEFKAGR 357
Db 298 IIDDTFDAYATFDELVTENDAIQRWDANAIDSIOQYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNCCKEFVKEFVRNLMVAKWANE-GHIPTTEHDPVVIITGGANLLTTTCVLGMSD 416
Db 358 LDRVYAKNEMKVLRAVFKETQMLNDCDHPKYEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IFTKESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKHSSSSLESYMEKYNNEEYA 476
Db 418 FISHETFEWMNSESIVRASALIPAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477

QY 477 QTLIYKEVDVWKDINREYLTTKNIPRLLMAVILCOFLEVOYAGKDNFTMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRPTVEPMFVLRLNLRVADTLYKEKDITYSTAKGLKXNM 537
QY 537 IKSLLVYPMISI 547
Db 538 INPILESVKI 548

RESULT 5
US-09-895-752-52
; Sequence 52, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-895-752-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;
QY 2 MALTEEKPIRPIANFPSPISWGDQFLIYKQVEQVQVNDLKEVRQLLKEALDIPMKH 61
Db 1 MAASADKCRPLANFHSVWGHFLSYTHETNQKEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETTYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNI FDASSKQNDNDNNLVVSLRFLVRQGHY 117
QY 118 VTCDVFNKYKNGAPKQSLANDVEGLELLELYEATSMRVPGEIILEDALGFTSRSLIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSLYEASHLRVNEEILEEALTFTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEPNLQOS 237
Db 178 NLSNNNSLKVEGEALTQPTMTLPRMGARKYISYIENNDHAHLLLKFAKLDENMLQK 237
QY 238 LHKEELSHVCXKWKAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT 297
Db 238 FHQRELSDLTRWKLDPANKYPYARDRLVECYFWILGVYFEPKYSRARKWMTKVLNLT 297
QY 298 LIDDTYDAYGTYEELKIITEAVERWSITCLDTPYMKPIYKLFMDITYTEMEEFKAGR 357
Db 298 IIDDTFDAYATFDELVTENDAIQRWDANAIDSIOQYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNCCKEFVKEFVRNLMVAKWANE-GHIPTTEHDPVVIITGGANLLTTTCVLGMSD 416
Db 358 LDRVYAKNEMKVLRAVFKETQMLNDCDHPKYEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IFTKESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKHSSSSLESYMEKYNNEEYA 476
Db 418 FISHETFEWMNSESIVRASALIPAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477

QY 477 QTLIYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVOYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRTEVPFMFLERVLNLRVADTLYKEKDTYSTAKGKLKNN 537
QY 537 IKSLIYVPMI 547
Db 538 INPILIESVKI 548

RESULT 6
US-09-903-012B-52
; Sequence 52, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6569656, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-903-012B-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;

QY 2 MALTEEKPIRPIANPPPSIWGDQFLIYQKQVEGVQIVNDLKEVROLLEALDIPMKH 61
Db 1 MAASSADKCRPLANFHPSPVWGVEYFSLYTHETNQEKVEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS-----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFIDASSKQNDNNLYVSLRFLVRQGHY 117
QY 118 VTCDVFNKYKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSYEASHLRVRNEEILEEALTTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEVGEALTQPIRWTLPRMGARKYISYIENNDAHHLLKFAKLDFFNLQK 237
QY 238 LHKEELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT 297
Db 238 FQRELSDLTRWKKLDLPANKYPYARDELVECYFWILGVYFEPKYSRARKMTKVLNLS 297
QY 298 LIDDTYDAYGTBYELKIFTEAVERWSITCLDTPYMKPIYKLFMDYTEMEEFLAKEGR 357
Db 298 IIDDTFDAYATFDELVTNDALQRWDANAIDSQPYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNGCKEVEFKVFNRLNVEAKWANE-GHPTTEHDPPVITGGANLLTTTCVLGMSD 416
Db 358 LDRVYAKNEMKULVRAFYKETQWLNDCDHI PKYEEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IFTKESVEMAVSAPPLFRYSGILGRLRLNDLMTKAEQERKHSSSLSLESYKENVNBEYA 476
Db 418 FISHEFEWLMNESVIVRASALIRAMNDIVGHEDEQERGHVASLIECYMKDYGASKOET 477

QY 477 QTLIYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVOYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRTEVPFMFLERVLNLRVADTLYKEKDTYSTAKGKLKNN 537
QY 537 IKSLIYVPMI 547
Db 538 INPILIESVKI 548

RESULT 7
US-09-900-797-52
; Sequence 52, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6645762, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-900-797-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;

QY 2 MALTEEKPIRPIANPPPSIWGDQFLIYQKQVEGVQIVNDLKEVROLLEALDIPMKH 61
Db 1 MAASSADKCRPLANFHPSPVWGVEYFSLYTHETNQEKVEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS-----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFIDASSKQNDNNLYVSLRFLVRQGHY 117
QY 118 VTCDVFNKYKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSYEASHLRVRNEEILEEALTTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEVGEALTQPIRWTLPRMGARKYISYIENNDAHHLLKFAKLDFFNLQK 237
QY 238 LHKEELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT 297
Db 238 FQRELSDLTRWKKLDLPANKYPYARDELVECYFWILGVYFEPKYSRARKMTKVLNLS 297
QY 298 LIDDTYDAYGTBYELKIFTEAVERWSITCLDTPYMKPIYKLFMDYTEMEEFLAKEGR 357
Db 298 IIDDTFDAYATFDELVTNDALQRWDANAIDSQPYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNGCKEVEFKVFNRLNVEAKWANE-GHPTTEHDPPVITGGANLLTTTCVLGMSD 416
Db 358 LDRVYAKNEMKULVRAFYKETQWLNDCDHI PKYEEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IFTKESVEMAVSAPPLFRYSGILGRLRLNDLMTKAEQERKHSSSLSLESYKENVNBEYA 476
Db 418 FISHEFEWLMNESVIVRASALIRAMNDIVGHEDEQERGHVASLIECYMKDYGASKOET 477
QY 477 QTLIYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVOYAGKDNFTRMGDEYKHL 536

Db 478 YIKFLKEVTNAWKDINKQFSRPTVEPMFVLNLTTRVADTLTKYKQDITYSTAKGKLKJNM 537
QY 537 IKSLVYVPMISI 547
Db 538 INPILIESVKI 548

RESULT 8

US-09-398-395A-32
; Sequence 32, Application US/09398395A
; Patent No. 8468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-398-395A-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;
QY 6 BEKIRPIANFPSPISWGQFLIY--OKVEQGVQVINDLKEVQLLKEALDIPMKHAN 63
Db 13 EEEIVRPIADPSLWGRDHFSLDNQIAGKYAQEIETLKEQSRIILSASSRRTL--AE 70
QY 64 LLKLIDEIQRIGIPVHFERIDHALQCIYET-----YGNWNGDRSSLSWFLMRKOG 115
Db 71 KLIDLIDIVERLGIAHFQKIDDDLDQYKADPNFEAHEYNDL---QTLVQFRLLRQHG 127
QY 116 YVTCDFVNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPEGEIILEDALGTRSLSIM 175
Db 128 YNISPGLFIRFQDAKGFKESLNCNDIKGLNLNLYEASHVTRTHGEDILEALAFSTAHLE-- 185
QY 176 TKDAFSTNPALFTETORALKQPLWKELPRIEAAQVIPPYQQQDSHNKTLKLAKLEFNLL 235
Db 186 -SAAPHLSKPSLQVTHALEQSLHKSIPRVETRYFISYEEEBEQKNDVLLQFAKLDNFLL 244
QY 236 QSLHKEELSHVCKWKWKAFDIKGNAPCLDRIVECYFWGLSGVGEPOYSRARVFFTKAVAV 295
Db 245 QMLHKQELSEVSRWKKDLDFVTLLPYARDRAVECFWTVGVAEPQYSQARVWLAKTIAM 304
QY 296 ITLDDTYDAYGTVEELKIFTEAVERWSITCLDTPLEYMKPIYKLFMDTYTEMEEFLEAKE 355
Db 305 ISIVDDTFDAYGIVKELEIYTDIQRWDISQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLFCNGCKEFVKEFVRNLNLYEAKWANEHGPTEEHDPVVIITGGANLLTTTCYLGMS 415
Db 365 GRSDVHVYAKEMKEIVRNVFVEAKWFTIEGYPVSEVLSNALATSTYLLTTTSYLGK 424
QY 416 DIFTKESVEWAVASAPPLFRYSGLIGRLNDLTHKAEQERKSSSLESYMKYNNVEEY 475
Db 425 SA-NKQDFEWLAKNPKILEANVTLCRVDDIATYVEKGRGQIATGIECMRDYGVSTEK 493
QY 476 AQTLTYKEVEDYWKDINREYLTTKNIPEPLMAVLYLCQFLEVQYA-GKDNFTRMGDEVK 534

Db 484 AMEKFOEMAETAWKQVNEGILRTPVSTETILTRILNLARIIDVTYKHNQDGYTHPEKVLK 543
QY 535 HLIKSLLVYPMISI 547
Db 544 PHIALLVDSIEI 556

RESULT 9

US-09-887-586A-32
; Sequence 32, Application US/09887586A
; Patent No. 6496354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64963541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-887-586A-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;
QY 6 BEKIRPIANFPSPISWGQFLIY--OKVEQGVQVINDLKEVQLLKEALDIPMKHAN 63
Db 13 EEEIVRPIADPSLWGRDHFSLDNQIAGKYAQEIETLKEQSRIILSASSRRTL--AE 70
QY 64 LLKLIDEIQRIGIPVHFERIDHALQCIYET-----YGNWNGDRSSLSWFLMRKOG 115
Db 71 KLIDLIDIVERLGIAHFQKIDDDLDQYKADPNFEAHEYNDL---QTLVQFRLLRQHG 127
QY 116 YVTCDFVNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPEGEIILEDALGTRSLSIM 175
Db 128 YNISPGLFIRFQDAKGFKESLNCNDIKGLNLNLYEASHVTRTHGEDILEALAFSTAHLE-- 185
QY 176 TKDAFSTNPALFTETORALKQPLWKELPRIEAAQVIPPYQQQDSHNKTLKLAKLEFNLL 235
Db 186 -SAAPHLSKPSLQVTHALEQSLHKSIPRVETRYFISYEEEBEQKNDVLLQFAKLDNFLL 244
QY 236 QSLHKEELSHVCKWKWKAFDIKGNAPCLDRIVECYFWGLSGVGEPOYSRARVFFTKAVAV 295
Db 245 QMLHKQELSEVSRWKKDLDFVTLLPYARDRAVECFWTVGVAEPQYSQARVWLAKTIAM 304
QY 296 ITLDDTYDAYGTVEELKIFTEAVERWSITCLDTPLEYMKPIYKLFMDTYTEMEEFLEAKE 355
Db 305 ISIVDDTFDAYGIVKELEIYTDIQRWDISQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLFCNGCKEFVKEFVRNLNLYEAKWANEHGPTEEHDPVVIITGGANLLTTTCYLGMS 415
Db 365 GRSDVHVYAKEMKEIVRNVFVEAKWFTIEGYPVSEVLSNALATSTYLLTTTSYLGK 424
QY 416 DIFTKESVEWAVASAPPLFRYSGLIGRLNDLTHKAEQERKSSSLESYMKYNNVEEY 475
Db 425 SA-NKQDFEWLAKNPKILEANVTLCRVDDIATYVEKGRGQIATGIECMRDYGVSTEK 483
QY 476 AQTLTYKEVEDYWKDINREYLTTKNIPEPLMAVLYLCQFLEVQYA-GKDNFTRMGDEVK 534
Db 484 AMEKFOEMAETAWKQVNEGILRTPVSTETILTRILNLARIIDVTYKHNQDGYTHPEKVLK 543

```
QY 535 HLIKSLLVYPM5I 547
Db 544 PHIALLVDSIEI 556

RESULT 10
US-09-895-752-32
; Sequence 32, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6559297, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-895-752-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 EEKPIRIANFPSSINGDOFLY--OKQVEQVEQIVNDLKEVRQLLKEALDIPMKHAN 63
Db 13 EEEIVRPIADPSPSLWGRFHSFLDNQIAGKYAQEIETLKEQSRILSASSRRL--AE 70
QY 64 LKLDIDEIQRIGIPHYPEREIDHALQCIYET-----YGDNMNGDRSSLWFLMRKQG 115
Db 71 KLDLIDIVERLGIAYHFEKQIDMDLQFYKADPNFEAHEYNDL---QTLVQVQFLLRQHG 127
QY 116 YVTCDFVNNYKNGAFKQSLANDVEGLLEYEATSMRVPGEIILEDALGFTRSLIM 175
Db 128 YNISPKLFIRFQDAKGFESLNCNDIKGLNLNLYEASHVRTHGEDILEEALAFSTAHL-- 185
QY 176 TKDAFSTNPALFTEIQRALKQPLMKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEFNL 235
Db 186 -SAAPHLKSPUSQVTHALEQSLKSIPTVETRYFISYEEEQKNDVLLQFALDFNLL 244
QY 236 QSLHKEELSHVCKWKAQPIKKNAPCLDRIVECYFVGLSGYEPQYSRARVFTKAVAV 295
Db 245 QMLHKQELSEVRWKKLDFTVTLPYARDRAVECYFTWVGVAEPQYSQARVMLAKTIAM 304
QY 296 ITLIDDTYDAYGTVEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEF 355
Db 305 ISIVDDTFDAYGIVKELEIYTDIQWDISQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLPNCGKEFVKEFVRNLMVBAKWANEHGIPTTEHDPPVITGGANLLTTTCYLGM 415
Db 365 GRSDVVHYAKERMEIVRNRYFVEAKFIEGMPVSEYLSNALATSYLLTTTSYLG 424
QY 416 DIFTKESVWAVSAPPLFRYSGLILRRNLDMTHKAEQERKSSSSLESYMKENYNEEY 475
Db 425 SA-NKQDFEWLAKNPXILEANVTLCRVDDIATYVEKGRGOIATGIECYMRDYG 483
QY 476 AQTLLIYKEVDWVDKINREYLTNKPRLMAVILCOFLEVOYA-GKDNFTRMGDEYK 534
Db 483 AQTLLIYKEVDWVDKINREYLTNKPRLMAVILCOFLEVOYA-GKDNFTRMGDEYK 534

RESULT 11
US-09-903-012B-32
; Sequence 32, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6569656, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-903-012B-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 EEKPIRIANFPSSINGDOFLY--OKQVEQVEQIVNDLKEVRQLLKEALDIPMKHAN 63
Db 13 EEEIVRPIADPSPSLWGRFHSFLDNQIAGKYAQEIETLKEQSRILSASSRRL--AE 70
QY 64 LKLDIDEIQRIGIPHYPEREIDHALQCIYET-----YGDNMNGDRSSLWFLMRKQG 115
Db 71 KLDLIDIVERLGIAYHFEKQIDMDLQFYKADPNFEAHEYNDL---QTLVQVQFLLRQHG 127
QY 116 YVTCDFVNNYKNGAFKQSLANDVEGLLEYEATSMRVPGEIILEDALGFTRSLIM 175
Db 128 YNISPKLFIRFQDAKGFESLNCNDIKGLNLNLYEASHVRTHGEDILEEALAFSTAHL-- 185
QY 176 TKDAFSTNPALFTEIQRALKQPLMKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEFNL 235
Db 186 -SAAPHLKSPUSQVTHALEQSLKSIPTVETRYFISYEEEQKNDVLLQFALDFNLL 244
QY 236 QSLHKEELSHVCKWKAQPIKKNAPCLDRIVECYFVGLSGYEPQYSRARVFTKAVAV 295
Db 245 QMLHKQELSEVRWKKLDFTVTLPYARDRAVECYFTWVGVAEPQYSQARVMLAKTIAM 304
QY 296 ITLIDDTYDAYGTVEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEF 355
Db 305 ISIVDDTFDAYGIVKELEIYTDIQWDISQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLPNCGKEFVKEFVRNLMVBAKWANEHGIPTTEHDPPVITGGANLLTTTCYLGM 415
Db 365 GRSDVVHYAKERMEIVRNRYFVEAKFIEGMPVSEYLSNALATSYLLTTTSYLG 424
QY 416 DIFTKESVWAVSAPPLFRYSGLILRRNLDMTHKAEQERKSSSSLESYMKENYNEEY 475
Db 425 SA-NKQDFEWLAKNPXILEANVTLCRVDDIATYVEKGRGOIATGIECYMRDYG 483
QY 476 AQTLLIYKEVDWVDKINREYLTNKPRLMAVILCOFLEVOYA-GKDNFTRMGDEYK 534
Db 483 AQTLLIYKEVDWVDKINREYLTNKPRLMAVILCOFLEVOYA-GKDNFTRMGDEYK 534
```

QY 535 HLIKSLVVPMSI 547
Db 544 PHIALLVDSIEI 556

RESULT 12

US-09-900-797-32
; Sequence 32, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 66457621, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-900-797-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 EEKRPPIANFPSSIWGDOFLIY--QKQVEQVEQIVNDLKKEVRQLLKEALDIPMKHAN 63
Db 13 EEEIVRPVADPSLWGDQFLSPSIKNQVAEIEALKQTRNML---LATGMKLAD 66
QY 64 LKLDIDEIQRIGIPYHFEREIDHALQCIYET-----YGDWNGDRSSIMFLMRKQG 115
Db 71 KLDLDIVERLGTAYHFEQIDMDLQFYKADPNFEAHEYNLD---QTLVQFLLRQHG 127
QY 116 YVYVCDVFNKYKNGAFKQSLANDVEGLELEYEATSMRVPGIIEBDALGFTSRSLSIM 175
Db 128 YNISPKLPIRFODAKGFKESLNDIKGLNLNLYEASHVRTHGEDILEEALAFSTAHL-- 185
QY 176 TKDAPSTNPALEIQRALKQPLWKRLPRIEAAQYIPFYQQDSHNKTLKLAKLEFN 235
Db 186 -SAAPHLKSPSQVTHALEQSLHSIPVETRYFISIEEEQKNDVLLQFALDFNLL 244
QY 236 QSLHKEELSHVCKWKAFDIKKNAPCLDRDRIVECYFWGLSGYEPQYSRARVFTKAV 295
Db 245 QMLHKEELSEVRWKDLDFVTTLPYARDRAVECYFWTMGVYAEQYSQARVMLAKTAM 304
QY 296 ITLIDTIDYDAYGYEELKIFTEAVERMSITCLDTPLEPMKPIYKLFWDITYTEMEEFLA 355
Db 305 ISIVDDTIDYDAYGIVKELEIYTDALQWIDISQIDRLPDYMKISYKALLDLYNDEMELSKD 364
QY 356 GRITDLNCGKEFKVEFVRNLMVEAKWANEHGIPTTEEDHPVVIITGGANLLTTTCYLGMS 415
Db 365 GRSDVVHYAKERMEKIVRNVFVEAKVFIETGMPVSEYLSNALATSTYLLTTTTSYLGMK 424
QY 416 DIFTKESVEMWASAPPLFRYSGILGRNLDMTHKAEQERKHSSESLSYMKENYNEEY 475
Db 425 SA-NKQDFWLAKNPKILEANVILCRVIDDIATYEVEKGGQIATGECWYRDYGVSTEK 483
QY 476 AQTLLIYKEVEDWVKDINREYLTNNIPRPLMAVILYLCQFLEVQYA--GKDNFTRMGDEYK 534
Db 484 AMEKFOEMAETAWKDYNEGILRTPPVSTETILTRILNARLAIIDVTYKKNQDGYTHPEKVK 543

QY 535 HLIKSLVVPMSI 547
Db 544 PHIALLVDSIEI 556

RESULT 13

US-09-398-395A-2
; Sequence 2, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-398-395A-2

Query Match 35.1%; Score 1020.5; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 2.3e-91;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;

QY 6 EEKRPPIANFPSSIWGDOFLIY--QKQVEQVEQIVNDLKKEVRQLLKEALDIPMKHAN 63
Db 10 EEEIVRPVADPSLWGDQFLSPSIKNQVAEIEALKQTRNML---LATGMKLAD 66
QY 64 LKLDIDEIQRIGIPYHFEREIDHALQCIYETVYGDWNGD-----RSSLWFLMRKQGYV 118
Db 67 TLNLDITIERLIGISYHFEKEIDLDQIY-----NQNSNCNDLCTLSALQFLLRQHGFI 121
QY 119 TCDVFNKYKNGAFKQSLANDVEGLELEYEATSMRVPGIIEBDALGFTSRSLSIMTKD 178
Db 122 SPEIFSKFDQNGKFKESLASDVLGLNLNLYEASHVRTHADDILEDALAFSTIHL-- 176
QY 179 AFTSNPALFT---EIQRALKQPLWKRLPRIEAAQVI-PFYQQDSHNKTLKLAKLEFN 233
Db 177 -SAAPHLKSPLEQVTHALEQCLHKGVPVTRTRFISIIYDKESKNNVLLRPAKLDFN 234
QY 234 LLOSLHKEELSHVCKWKAFDIKKNAPCLDRDRIVECYFWGLSGYEPQYSRARVFTKAV 293
Db 235 LLOMLHKEELAQVSRWKDLDFVTTLPYARDRAVECYFWALGVYFEPQYSQARVMLVKT 294
QY 294 AVITLIDTIDYDAYGYEELKIFTEAVERMSITCLDTPLEPMKPIYKLFWDITYTEMEEFLA 353
Db 295 SMISIVDDTIDYDAYGIVKELEIYTDALQWIDINREYLTNNIPRPLMAVILYLCQFLEVQY 354
QY 354 KEGRTDLNCGKEFKVEFVRNLMVEAKWANEHGIPTTEEDHPVVIITGGANLLTTTCYLG 413
Db 355 SAGSHIYCHAIERKKEVVRNVRNVSTFIEGTTPVSEYLSNALATTTTYLLATTSYLG 414
QY 414 MSDIFTKESVEMWASAPPLFRYSGILGRNLDMTHKAEQERKHSSESLSYMKENYNEEY 473
Db 415 MKSA-TEQDFEFLSKNPKILEASVILCRVIDDTATVEVEKSRQIATGIECCNRDYGIST 473
QY 474 EYATLLIYKEVEDWVKDINREYLTNNIPRPLMAVILYLCQFLEVQYA--GKDNFTRMGDE 532
Db 474 KEAMAKFOEMAETAWKDINREYLTNNIPRPLMAVILYLCQFLEVQYA--GKDNFTRMGDE 532
QY 533 YKHLIKSLVVPMSI 547

```
Db 534 LKPHIINLLVDSIKI 548

RESULT 14
US-09-887-586A-2
; Sequence 2, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-887-586A-2

Query Match 35.1%; Score 1020.5; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 2.3e-91;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;

Qy 6 BEKPIRPIANPPPSIWGDQFLY--QKQVEQGVQIVNDLKEVQRLKKEALDIPMKHAN 63
Db 10 EEEIVRPVADFPSPSLWGDQFLSPSIKNQVAEYQAEIALKEQTRNML---LATGMKLD 66
Qy 64 LKLDLDELQIGPHFPERIDHALOCIVETYGDNNGD-----RSSLWFLMRKQGVV 118
Db 67 TLNLDITIERLGISYHFEKIDIDLOIY-----NQNNSCNDICTSALQFLLRQHGFI 121
Qy 119 TCDVFNKYKQNGAFKQSLANDVEGLELYEATSMRVPGIILEDALGFRSRLSINTKD 178
Db 122 SPEIFSKFODENGKFKESLASDVLGLNLYEASHVTHADDDILEDALAFSTHLE----- 176
Qy 179 AFTSNPALFT---EIQRALKQPLWKRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFIFSIYDKESQSKNNVLLRFAKLDFN 234
Qy 234 LLOSLHKEELSHVCKMKKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRABVFTKAV 293
Db 235 LLOMLHKQELAQVSRWKKDLDFVTTLFYARDRVVECYFWALGVYFEPQYSQARVMLVXTI 294
Qy 294 AVITLDDTYDAYGYEELKIFTEAVERWSITCLDTPYMKPIYKLFMDTYTEMEBEFLA 353
Db 295 SMSIVDDTFDAYGTVEKLEAYTDAIQRDWINEIDRLPDYMKISYKAILDLYKDYKELS 354
Qy 354 KEGRTDLFNCGEFVKFEFVRNLMVEAKWANEHIPPTEEHDPVVIITGGANLLTTTCYLG 413
Db 355 SAGRSHIVCHATERMKEVRNVNVESTWTFIEGYTPPVSEVLSNALATTTTYYLATTSYLG 414
Qy 414 MSDIFTKESVEMAVSAPPLFRYSGILGRRLNDLMTKAEQKHSSESSLESYKENVNNE 473
Db 415 MKSA-TEQDFEWSLKNPKILEASVIIICRVDDTATYVEKSRGQIATGECNRDYGIST 473
Qy 474 EYAQTLIYKEVEDVWKDINREYLTTKNIPRPLMAVILYLCQFLEVOYA-GKDNFTRMGDE 532
Db 474 KEAMAKFQMAETAWKDINEGLRPTVPVSTEFILNLARIVEVYIHNLDGTHPEKV 533
Qy 533 YKHLIKSLLYVPMISI 547

Query Match 35.1%; Score 1020.5; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 2.3e-91;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;

Qy 6 BEKPIRPIANPPPSIWGDQFLY--QKQVEQGVQIVNDLKEVQRLKKEALDIPMKHAN 63
Db 10 EEEIVRPVADFPSPSLWGDQFLSPSIKNQVAEYQAEIALKEQTRNML---LATGMKLD 66
Qy 64 LKLDLDELQIGPHFPERIDHALOCIVETYGDNNGD-----RSSLWFLMRKQGVV 118
Db 67 TLNLDITIERLGISYHFEKIDIDLOIY-----NQNNSCNDICTSALQFLLRQHGFI 121
Qy 119 TCDVFNKYKQNGAFKQSLANDVEGLELYEATSMRVPGIILEDALGFRSRLSINTKD 178
Db 122 SPEIFSKFODENGKFKESLASDVLGLNLYEASHVTHADDDILEDALAFSTHLE----- 176
Qy 179 AFTSNPALFT---EIQRALKQPLWKRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFIFSIYDKESQSKNNVLLRFAKLDFN 234
Qy 234 LLOSLHKEELSHVCKMKKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRABVFTKAV 293
Db 235 LLOMLHKQELAQVSRWKKDLDFVTTLFYARDRVVECYFWALGVYFEPQYSQARVMLVXTI 294
Qy 294 AVITLDDTYDAYGYEELKIFTEAVERWSITCLDTPYMKPIYKLFMDTYTEMEBEFLA 353
Db 295 SMSIVDDTFDAYGTVEKLEAYTDAIQRDWINEIDRLPDYMKISYKAILDLYKDYKELS 354
Qy 354 KEGRTDLFNCGEFVKFEFVRNLMVEAKWANEHIPPTEEHDPVVIITGGANLLTTTCYLG 413
Db 355 SAGRSHIVCHATERMKEVRNVNVESTWTFIEGYTPPVSEVLSNALATTTTYYLATTSYLG 414
Qy 414 MSDIFTKESVEMAVSAPPLFRYSGILGRRLNDLMTKAEQKHSSESSLESYKENVNNE 473
Db 415 MKSA-TEQDFEWSLKNPKILEASVIIICRVDDTATYVEKSRGQIATGECNRDYGIST 473
Qy 474 EYAQTLIYKEVEDVWKDINREYLTTKNIPRPLMAVILYLCQFLEVOYA-GKDNFTRMGDE 532
Db 474 KEAMAKFQMAETAWKDINEGLRPTVPVSTEFILNLARIVEVYIHNLDGTHPEKV 533
Qy 533 YKHLIKSLLYVPMISI 547
```

Db 534 LKPHIINLLVDSIKI 548

Search completed: June 3, 2004, 16:34:05
Job time : 31 secs

121	VFNNYKDNGAFKQSLANDVEGLELEVEATSMRVPGEIILEDALGFTRSRLSINTKDAFS	180
182	TNPALEIQRALKQPLWKELPRIEAAQYIPFYQQDSHNKTLKLAKLEFNLIQSLHKE	241
181	TNPALEIQRALKQPLWKELPRIEAAQYIPFYQQDSHNKTLKLAKLEFNLIQSLHKE	240
242	ELSHVCKWKAFAPIKKNAPCLDRIVECYFWGLSGGYEPOYSRARVFTKAVAVITLIDD	301
241	ELSHVCKWKAFAPIKKNAPCLDRIVECYFWGLSGGYEPOYSRARVFTKAVAVITLIDD	300
302	TYDAYGTYEELKFTTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEEFKLAKGRDOLF	361
301	TYDAYGTYEELKFTTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEEFKLAKGRDOLF	360
362	NCCKEFVKEFVRNLMVEAKWANEGHIPPTBEHDPWIIITGGANLLTTTCYLGMSDITFKE	421
361	NCCKEFVKEFVRNLMVEAKWANEGHIPPTBEHDPWIIITGGANLLTTTCYLGMSDITFKE	420
422	SVEWAVSAPPLFRYSGILGRRLNDLMTWKAEQERKHSSSLESYMKYNNVEEYAQTLIY	481
421	SVEWAVSAPPLFRYSGILGRRLNDLMTWKAEQERKHSSSLESYMKYNNVEEYAQTLIY	480
482	KEVEDYWKDINRBYLTTKNIPRPLLMAVIVLCQPLEVQYQAKGNFTRMGDEYKHLIKSLL	541
481	KEVEDYWKDINRBYLTTKNIPRPLLMAVIVLCQPLEVQYQAKGNFTRMGDEYKHLIKSLL	540
542	VYPMSTGS	549
541	VYPMSTGS	548

```

RESULT 2
US-10-411-066-40
; Sequence 40, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 551
; TYPE: PR1
; ORGANISM: Artemisia annua
US-10-411-066-40

```

Query Match	50.3%	Score 1462.5	DB 15	Length 551
Best Local Similarity	52.0%	Pred. No. 1.2e+126		
Matches 282	Conservative 97	Mismatches 162	Indels 1	Gaps 1
2	MALTEEKPIRPANFPFSGWGDQFLIYQKQVEQGVQIVNDLKEVQRQLIKKALDIPMKH	61		
	1	MSLIVEDVRPNANPSEIWDQFLAYDQDEQEGVEQVIKDLKEEVKSELLTALNSPTQH	60	
	62	ANLLKLIDIEQRLGIPYHFEREDHALQCIYETVYGNWNGDRSLFRLMKRGYVYTC	121	
	61	TELLKFIDALERLGIAYVFEESINQVFQHYTAYGDKWTGNTSLFRLNRQIGFFVSSD	120	
	122	VNNYKDKNGAKQSLANDVEGLLELYEATSRVPGEIILEDALGTRSRSLSTMTKDAFS	181	
	121	ISTYNDKEGRPKESLEKOVGHGLELYEAAFMVPEGILDDALVTRTCLDEIAKNPSL	180	
	162	TNPALFTETIQRALQKPLWKRLPRIEAAQVIPPYQQDDSHNKTLLKLAKLEFNLLQSLHKE	241	

	Query Match	38.6%;	Score 1123;	DB 9;	Length 548;
	Best Local Similarity	38.8%;	Pred. No. 4e-95;		
	Matches 214;	Conservative 123;	Mismatches 206;	Indels 8;	Gaps 3
QY	2	MALTEEKPIRPIANPPPSIGWDQFIYIKQVQEQVEQVINDLKCEVRQLLKEALDIPMKH	61		
DB	1	MAASSADKCRPLANPHPSVWGVIHFLSYTHEITNQEKVEVDEYKETIRKMLVETCD	57		
QY	62	ANLLKLDEIORLGIPIYPHFREIDHALOCIXITYGDNWNGDRS	117		
DB	58	TQKLVLDAMQRLGVAYHFDNEIFTSIQNIFDASSKQNDNNLVSVLRFRLVRQGHY	117		
QY	118	VTCDVFNNYKDKNGAFKOSLANDVEGLLELYEATSMRYPGCEIILEDALGFTSRSLSIMTK	177		
DB	118	MSSDVFKQFTNQDGKFKETLVNDVQGLSLYKASHLRVRNEEILBEALFTTTHUESIVS	177		
QY	178	DAFSTNPALFTEIORALQKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQS	237		

Query Match	38.6%	Score 1123	DB 10	Length 548
Best Local Similarity	38.8%	Pred. No. 4e-95		
Matches 214	Conservative 123	Mismatches 206	Indels 8	Gaps 3
Qy	2	MALTEEKPIRPIANFPFSIWGDQFLIYKQVQGVQIVNDLKKEVRQLLKALDIPMKH	61	
Db	1	MAASSADKCRPLANFHSFVWGHHFLSYTHITNQEKVEVDYKETIRKMLVETCD--NS	57	
Qy	62	ANLLKLIDEIQRIGIPYHPREBDHALQCIYETVYGDWNGDRS-----SLWFLMRKQGY	117	
Db	58	TQKLVLDAMQRLGVAYHFNENIETSIQNTFPDASSKQNDNDNLVYVSLRFLRVQOQGHY	117	
Qy	118	VTCDVFNNYKDKGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGTRSRSLSMTK	177	
Db	118	MSSDVPKQFTNQDGKFKETITNDVQGLLSLYEASHLAVRNEEILEEALFTTTTHLESIVS	177	
Qy	178	DARSTNPALFTEIQRALKQPLWKLPRLEIAQVIPPYQQQDSHNKTKILKLAKLEFNILLOS	237	
Db	178	NLGNNSNLKVEYGEALTQPIRTPMPMGARKYIISIVENDAHHLLKLFKALQDFNMLQK	237	

Qy	238	LHKEELSHVCKWKWAFDIKKNAPCLDRDRIVECYFGLSGYBPQYSRARVFFTKA VAVIT	297
Dd	238	FHQRELSLTRWKKDLDPANKYPYARDLVECYFWILGVYFPKYSRARKMWTKVNLITS	297
Qy	298	LIDDTYDAYGYBYBLIFTEAVERWSIICLDTLPEYMKPIYLKFMDTYTEMEEFLAKEGR	357
Dd	298	IIDTTFAYAYFDELVTNDAIQRWDANAISIQYMPRPAQALJDIYSEMQVLSKEGK	357
Qy	358	TDLFNCCGEFKFYKFVRNLMWAKWANE-CHIPTTBEEHPDVIIITOGANLLTTTCVLGMED	416
Dd	358	LDRVYAKVEMKULVRAFYKETOWMLNDCDHIPKYEEQVENAIVSAGYMMISITCVLGISE	417
Qy	417	IFTKESVEWNAVSAPPLFRYSIGILGERLNDLMTHKABOERKHSSSLSESVMKYNVNNEVA	476
Dd	418	FISHETTEWLNMNESVIWAASALIARAMNDIVGHDEQEGRGHVASLTCEYMKDYGASKOET	477
Qy	477	QTLLTYKEVEDVVKDKINREYLTTONIPRLMAVIVLCOFLEVQYAKGNFTRMGDYEKHL	536
Dd	478	YIKELKEVTNAWKDKINKOPFSRPTVEPMFVLBRVLNTAVADTLTYKEKOTIYSTAKGLKNM	537
Qy	537	IKSLLVTPMSI	547
Dd	538	INPILIESVKI	548

RESULT 6

```

US-09-893-820-52
; Sequence 52, Application US/09893820
; Publication No. US2004005386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US2004005386A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-893-820-52

```

Query Match	38.6%;	Score 1123;	DB 12;	Length 548;
Best Local Similarity	38.8%;	Pred. No. 4e-95;		
Matches 214;	Conservative 123;	Mismatches 206;	Indels 8;	Gaps 3
QY	2	MALTEKPIRPIANPPPSINGDQFIYIKQVQGVQEIIVNDLKXEVROLLKEALDIPMKH	61	
Db	1	MASSADKCRPLANFHPSPVWGVIHFSYTHEITINQEKVEVDEYKETIRKMLVETCD--NS	57	
QY	62	ANLLKLDEIORLGTPYHFEREIDHALOCIVETYGDNWNGDRS---SLWFLMRKQGY	117	
Db	58	TQKLVLIDAMQRLGVAFHFDNEIETISQNIIPASSKQNDNNLVVSLRFLRVQGGHY	117	
QY	118	VTDVFNMYKQNGAFKOSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLIMTK	177	
Db	118	MSSVDFKQFTNODGKFRETLINDVGLISLYEASHLRVRNEIILBEALFTTTHUESIVS	177	
QY	178	DAFSTNPALFTEIORALKQPLMKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEPNLLOS	237	

178	NLSNNNSLKVEGEALTOPTRMTLPWGARKYISIEYNNDAAHHLLLPKAKLDFNNLOK	237
Dd		
238	LHKEELSHVCKWKAFAFIKKAPCLDRDIVECYFWGLSGGYBPQYSRARVFFTKAVAVIT	297
Qy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Dd		
238	PHQESDLTLRWKLDLFANKYPYARDLVCEYFWILGVFPFKYSRAIKMWTKVNLNIS	297
Dd		
298	LIDDTDAYGTYEELKIIFTEAVERWSIITCLDTPLEYMKPIYKLFMDTYTMEEFFLAKEGR	357
Qy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Dd		
298	IIDDTFAYATFDELVTENDAIQRWDANAIDSIQPYMRPAYQALLDIYSEMEQVLSKEGK	357
Dd		
358	TDLFNCGCFEYKFVRNLMWEAKWANE-CHPTTTEEHDPVLIITGGANLLTTTCYLGMSD	416
Qy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Dd		
358	LDRVYAKNEVKKLVRAYFKETQMDCDHIPKYEEQVENAIVSAGYMMISTTCLVGISE	417
Dd		
417	IFTXESVEWAYSAAPLPYSGILGRRLDLMTHKAEQERKHSSSILESYMKZYNNVNEEYA	476
Qy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Dd		
418	FISHETTEWLNMNESVIVASALIAAMNDIVGHEDEQERGHVASILECYMKDYGASKOET	477
Dd		
477	QTLIYKEVEDVMKDINREBLATTQNIPRELLMAVIYLCOFLEVQYAGKNFTRMGDYBYKHL	536
Qy	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Dd		
478	YIKFLKEVTNAWKDINKQFSRPTPEVMPFVLERVNLTRVADTLYEKCDTYSTAKGKLKMM	537
Dd		
537	IKSLLVTPMSI 547	Qy
Dd	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
538	INPILIESVKI 548	Dd

RESIT.T 7

```

US-10-424-599-186274
; Sequence 186274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285584
; SEQ ID NO 186274
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_i39218C.1.pep
US-10-424-599-186274

```

Query Match	37.8%	Score 1099.5;	DB 12;	Length 567;
Best Local Similarity	40.5%;	Pred. No. 6.4e-93;		
Matches 227:	Conservative	115;	Mismatches 176;	Indels 43;
	Gaps	10;		

Qy	10	IRPIANPPPSIWGQDFLIYQKQVEQIVND-----LKKEVRQLLKEALDIPMK	60
Db	22	IRNTANTSPSPVMGDFLYX-----VPSSVEDDSHIKQAOULTKCEVRMLTAPID----	70
Qy	61	HANL--LKLDETORLGIPIHFEEREIDHALQCIYE-----TYGDNWNGDRSSLWF	108
Db	71	NNFYFKLEUFDTSVQRGLGVSHFEHEIDGALHQLIYNISTKONNIITHDDDL-----CHVALLF	127
Qy	109	RIIMEKQGYVYTCDFNNYKDXNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGET	168
Db	128	RLLRQQGYHISSTNYFYFKDQTRNFSKAAANDIQGMLSLYEAAELRMHGEDIIEEAHNEA	187
Qy	169	RSRL--SMTKDASTNPALTEORALKQPLWKELPRIEAAQYIPFYQQQDSHNKTLK	226
Db	188	LVQUTKSLTQ-----LSPSMTAQVHSURRSURKGLPREATYMSFVEEDSSHDEKLT	243
Qy	227	LAKLEFNLQLSLHKEEASHVCKW--KAPDIKNAFCLDRIVECYFWGLGSGYEPQYSRA	285
Db	244	FAKLDENMLQELHOKEVNNVTRWIKLNVSTKLPFVEDRTIAECYFWILGIFYEPQYSIA	303

QY		296	ITLIDDTDAYGTVEELKIFTEAVERMSITCLDTPLEYNKPTIKLFMDTYTMEBEFLAKE	355
Db		305	ISIVDDTFDAYGIVKELEIYTDAIQRWDISOIDRLPDYMKISYKALLDYNDYEMELSKD	364
QY		356	GRTDLFCNGKEFVZEFTRNLNVEAKWANEHGIPTTEBHDVVUITGGANLLTTTCYLGMWS	415
Db		365	GRSDVHYAKRMKEIVRNTVFVEAKWETEGYMPVPVSEYLSNALASTYTLTTTTSYLGMK	424
QY		416	DITKESVEVAVSAPPLFRYSGLIGRRLANDIMTHKAQEERKHSSSLSLYMKEYNNBYEY	475
Db		425	SA-NKQDFEWLANPKILEANNVTICRVDDIATYEVEGRQIATGIBCYMRDYGUSTEK	483
QY		476	AQTLIYKEVEDVMKDINREYLTNTKNIPRLMAVIYLCQFLEVQYA-GKDNFTRMGDYVK	534
Db		484	AMEKFQWEAETAWKVDNEGILRPPTVSTFEILTRIINLAARIIDVTYKHNQDGYTHEPKVLK	543
QY		535	HLIKSLILYPMSI	547
Db		544	PHIALLVDSIEI	556

```

RESULT 10
US-09-900-797-32
; Sequence 32, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-900-797-32

```

[illegible]

Query Match	35.8%	Score 1041	DB 12	Length 556
Best Local Similarity	39.6%	Pred. No. 1.7e-87		
Matches 219	Conservative 105	Mismatches 209	Indels 20	Gaps 7

Qy	6	E E K P I R P I A N F P S I W G D Q L Y I - - Q K V E Q B G V E Q I V N D L K V E Q R L L K E A L D I P M K H A N	63
D b	13	E E B I R P I A D P S L W G D R F H S F S L O N Q I A G K V A Q E I T L K E O S R I I L S A S S R T L - - A E	70
Qy	64	L L K L I D E I Q R L G P Y H F E R I D H A L Q C I Y E T - - - - - Y G D N W N G D R S S I W P R L M R K O G	115
D b	71	K L D L I D I V E R L G I A Y H F E K Q I D M L D O F Y K A D P N F E A H E Y N D L - - - - - Q T L S V O F R L L Q H G	127
Qy	116	Y Y V T C D V F N N Y K D K N G A F K O S L A N D V G L L E J Y E A T S M R V P G E I I L S D A L G F T T S R I S I M	175
D b	128	Y N I - S P K L F I R P Q D A K G F K E S L C N D I K G L L N L Y E A S H V R T H G E D I L E A A F S T A H L E - -	185
Qy	176	T K O A F S T N P A L F T E I Q R A L K O P L W K R I P R T E A A Q Y I P F Y Q O O S H N K T L L K A L B E N L L	235
D b	186	- S A A P H U K S P L S K Q V T H A L S Q S L H K S I P R V E T R Y F I S I Y E E E Q K D V L Q F A K L D E N L L	244
Qy	236	Q S L K E E L S H V C K W K A F D I K C N A P C L U R D R I V E C Y F W G L O S G V E P P Q Y S R A R V P T T K A V A V	295
D b	245	O M L F K O S L E V S R W K D L D F T T T L P Y A R D R A V E C Y F T M G Y A P F O S Q A R V M L A K T I M	304

Qy 296 ITLIDDTVDAGTYBELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDYTEMEEFLAKE 355
Db 305 ISIVDDTDFDAYGIVKELEYTDIAIQRWDISQIDRLPDYMKISYKALLDLNDYEMELSKD 364
Qy 356 GRITLFCNGKRFVKEFVRNLVWEAKWANEHGHIPTEEHDPVVIITGGANLTTTCYLGMS 415
Db 365 GRSDVVHAKERKKEIVRNFYFEAKWFIEGYPPVSEYLSNALATYLLTTSYLGMK 424
Qy 416 DIFTKESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSHSSSSLESYMKENYNEEY 475
Db 425 SA-NKQDFEWLAKNPKILEANVTLCRVIDDIAITYEVEKGRGQIATGIECYMRDYGSTEK 483
Qy 476 AQTLLYKEVEDVKDINREYLTKNIPRLLMAVIVLCQFLEVQYA-GKDNFTRMGDEYK 534
Db 484 AMEKFOEATKAWKQNEGILRPTPVSTBILTRILNARIIDVTYKHNDGYTHPEKVLK 543
Qy 535 HLIKSLLVYPMXI 547
Db 544 PHIIALLVDSIEI 556

RESULT 12
US-09-887-586A-2
; Sequence 2, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2003-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-887-586A-2

Query Match 35.1%; Score 1020.5; DB 9; Length 548;
Best Local Similarity 40.0%; Pred. No. 1.3e-85;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;
Qy 6 BEKPIRPIANPPSPISWGDQFLIY--QKVEQGVQIIVNDLKEVQRLKALDIPMKHAN 63
Db 10 EEEIYRPVADFSPLWGDQFLSFKIKQVAKYAEIALKQETRNML---LATGMKLAD 66
Qy 64 LLKLIIDEIQRIGIPYHFEREIDHALQCIYETVGDWNGD-----RSSLWFLMRKQGYV 118
Db 67 TLNLDITRIGISYHFEKEIDDLQIY-----NONSNCNDLCTSGALQFRLRQHGFI 121
Qy 119 TCDVFNNYKNGAFKQSLANDVEGLLELYEATSMRVPGEEIILEDALGFTSRSLSTMKD 178
Db 122 SPEIFSKFDENGKFKESLASDVLGLNLVEASHVTRHADDILEDALAFSTIHLE----- 176
Qy 179 AFTSNPALFT----EIQRALKQPLWRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHGKGVPRVETRFPISSIIDYKESQKNNVLLRFAKLDFN 234
Qy 234 LLQSLHKEELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAV 293
Db 235 LLQMLHKEELAQVSRWKKDLDFTVTLPYARDRVVECYFWALGVYFEPQYSQARVMLVKTI 294

Qy 294 AVITLIDDTVDAGTYBELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDYTEMEEFLA 353
Db 295 SMISIVDDTDFDAYGIVKELEYTDIAIQRWDINEIDRLPDYMKISYKAILDLNDYKEKELS 354
Qy 354 KEGRTDLFCNGKRFVKEFVRNLVWEAKWANEHGHIPTEEHDPVVIITGGANLTTTCYLG 413
Db 355 SAGSHIVCHAIERKKEIVRNFYFEAGYPPVSEYLSNALATYLLTTSYLG 414
Qy 414 MSDIFTKESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSHSSSSLESYMKENYNE 473
Db 415 MKSA-TEQDFEWLAKNPKILEASVTCRVIDDIAITYEVEKSRGQIATGIECCMRDYGIST 473
Qy 474 EYAQTLLYKEVEDVKDINREYLTKNIPRLLMAVIVLCQFLEVQYA-GKDNFTRMGDE 532
Db 474 KEAWAKFQNEATKAWKQNEGILRPTPVSTBILTRILNARIIVEVTYIHNLDGYTHPEKV 533
Qy 533 YKHLIKSLLVYPMXI 547
Db 534 LKPHIINLLVDSIKI 548

RESULT 13
US-09-903-012-2
; Sequence 2, Application US/099030312
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-903-012-2

Query Match 35.1%; Score 1020.5; DB 9; Length 548;
Best Local Similarity 40.0%; Pred. No. 1.3e-85;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;
Qy 6 BEKPIRPIANPPSPISWGDQFLIY--QKVEQGVQIIVNDLKEVQRLKALDIPMKHAN 63
Db 10 EEEIYRPVADFSPLWGDQFLSFKIKQVAKYAEIALKQETRNML---LATGMKLAD 66
Qy 64 LLKLIIDEIQRIGIPYHFEREIDHALQCIYETVGDWNGD-----RSSLWFLMRKQGYV 118
Db 67 TLNLDITRIGISYHFEKEIDDLQIY-----NONSNCNDLCTSGALQFRLRQHGFI 121
Qy 119 TCDVFNNYKNGAFKQSLANDVEGLLELYEATSMRVPGEEIILEDALGFTSRSLSTMKD 178
Db 122 SPEIFSKFDENGKFKESLASDVLGLNLVEASHVTRHADDILEDALAFSTIHLE----- 176
Qy 179 AFTSNPALFT----EIQRALKQPLWRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHGKGVPRVETRFPISSIIDYKESQKNNVLLRFAKLDFN 234
Qy 234 LLQSLHKEELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAV 293
Db 235 LLQMLHKEELAQVSRWKKDLDFTVTLPYARDRVVECYFWALGVYFEPQYSQARVMLVKTI 294

Qy	294	AVITLIDDTYDAYGYBELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEEFIA	353
Db	295	SMISIVDDTFDAYGTVKELEAYTDAIQWDINEIDRLPDYMKISYKAILDLYKDEKELS	354
Qy	354	KEGRDILFNCGKEFVKEFVRNLMVEAKWANECHIFTEEHDPVWITGGANLITTCYLG	413
Db	355	SAGSHIVCHAIERNKEVVRNINVESTWIEGYTPVSEYLSNALATTTYYLATTSYLG	414
Qy	414	MSDIFTKESVEWAVSAPPLFRYSGLRRRLNDLMTKAEQERKHSSSLESYMKYNNVNE	473
Db	415	MKSA-TEQDPFWLSKNPKILEASVILICRVIDDTATYEVEKSRGQIATGIECCMRDYGIST	473
Qy	474	EYAGTLIYKEVEDVWKDINREYLLTKNIPRPLMAVIVLCQFLEYQYA-GKDNFTRMGDE	532
Db	474	KEAMAKFQNMATKWDINEGLLRPTVPSTBELTILNLARIVEVTYIHNLDGYTHPEKV	533
Qy	533	YKHLIKSLVYPMISI	547
Db	534	LKPHIINLLVDSIKI	548

Search completed: June 3, 2004, 16:38:54
Job time : 52 secs

Db 21 IRPKADFPQSIWGDFFLPCDKNIDAGTEKRHOOLKEVRKMIV-----APMANSTOKLAF 76
Qy 68 IDEIQRLGIPYHFERIDHALQCIYETGDNWNG--DRSSILWFLMRKQGYVYTCDFVNNY 126
Db 77 IDSVQRLGVSHFTKEIEDELENIYHNNDANDLVTSTLRFELLREHGYNVSCDVFNKF 136
Qy 127 KDKNGAFQKSLANDVEGELLELYEATSMRVPGEIILEDALGFTSRSLSIMTKDAFSTNPAL 186
Db 137 KDEQGNFKSSVTSVQGLLELYQASLYRVHGEDIIDEAISFTTNHLSLAVS---SLDHPL 193
Qy 187 FTEIQRLKQPLWKLPRTEAAQYIPFYQQQDSHNKTLKLALEFNLLQSLHKEELSHV 246
Db 194 SEEVSHALKQSIIRGLPRVEARHYSVYQDIESHNKALLEFAKIDFNMLQFLHRELSEI 253
Qy 247 CKWKAFDIKKNAPCLDRIVECYFGLSGVGEPOYSRVRVFTTKAVAVITLDDTYDAY 306
Db 254 CRWKKDLDFQRLPYARDRVVGVFWISGVYFEPQYSLGRKMLTKVIAMASIVDDTYDSY 313
Qy 307 GTYEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTMEEFLLAKEGRDILFNGKE 366
Db 314 ATYEELIPYTNARWDIKCIDPELPEYMKPSYKALLDVYKEMEQSLVAEHGROYRVEYAKN 373
Qy 367 FVKEFVNLVMEAKWANEHGIPTTEBHDVPVITGGANLLTTTCYLGMSDIFTKESVEWA 426
Db 374 AMIRLAQSVLVEARWTLQNYKSFEEFKANALPTCGYAMLATISFVGMGDIPTPEFKWA 433
Qy 427 VSAPPLFRYSIGILGRRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEVAQTLIYKEVED 486
Db 434 ANDPKIIQASTIICRFMDVVAEHKFKHRRDDCSAIECYMEYGVSAQAEYDVFNKHVES 493
Qy 487 VKKDINREYLTNNIPRPLMAVILYQCLEVOYAGKDNFTMGDEYKHILKSLIYVPM 546
Db 494 AWKDNQEFKQKTEPTEVNLRLNRLARVMDVLYREGDGYTYVYVGAAGKGTISLLIEPIA 553
Qy 547 I 547
Db 554 L 554

RESULT 2

DCS1 GOSAR
ID DCS1 GOSAR STANDARD; PRT; 554 AA.
AC Q39761;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE (+)-delta-cadinene synthase isozyme XC1 (EC 4.2.3.13) (D-cadinene synthase).
OS Gossypium arboreum (Tree cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=29729;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nanking.
RX MEDLINE=96132653; PubMed=8554317;
RA Chen X.-Y., Chen X., Heinstein P., Davissou V.J.;
RA "Cloning, expression, and characterization of (+)-delta-cadinene
RT synthase: a catalyst for cotton phytoalexin biosynthesis."
RL Arch. Biochem. Biophys. 324:255-266(1995).
CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC diphosphate (FPP) to (+)-delta-cadinene.
CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC delta-cadinene + diphosphate.
CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC first (committed) step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U23206; AAA93064.1; -.
DR PIR; S68365; S68365.
DR HSP; Q40577; SEAU.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR Pfam; PF01397; Terpene synth; I.
DR Pfam; PF03936; Terpene synth_C; 1.
KW Lyase; Multigene family.
FT ACT_SITE 451 BY SIMILARITY.
FT ACT_SITE 527 BY SIMILARITY.
FT ACT_SITE 531 BY SIMILARITY.
SQ SEQUENCE 554 AA; 64137 MW; 59D6922DEDF9DCAF CRC64;

Query Match 39.3%; Score 1143; DB 1; Length 554;

Best Local Similarity 42.1%; Pred. No. 2.3e-70;
Matches 228; Conservative 106; Mismatches 197; Indels 10; Gaps 5;

Qy 10 IRPIANPPSIWGDFFL-IYOKQVEQGVQIWNLDKKEVRQLLKEALDIPMKHA-NLLKL 67
Db 21 MRPKADFPQSIWGDFFLPCDKNIDAGTEKRHOOLKEVRKMIV-----APMANSTOKLAF 76
Qy 68 IDEIQRLGIPYHFERIDHALQCIYETGDNWNG--DRSSILWFLMRKQGYVYTCDFVNNY 126
Db 77 IDSVQRLGVSHFTKEIEDELENIYHNNDANDLVTSTLRFELLREHGYNVSCDVFNKF 136
Qy 127 KDKNGAFQKSLANDVEGELLELYEATSMRVPGEIILEDALGFTSRSLSIMTKDAFSTNPAL 186
Db 137 KDEQGNFKSSVTSVQGLLELYQASLYRVHGEDIIDEAISFTTNHLSLAVS---SLDHPL 193
Qy 187 FTEIQRLKQPLWKLPRTEAAQYIPFYQQQDSHNKTLKLALEFNLLQSLHKEELSHV 246
Db 194 SEEVSHALKQSIIRGLPRVEARHYSVYQDIESHNKALLEFAKIDFNMLQFLHRELSEI 253
Qy 247 CKWKAFDIKKNAPCLDRIVECYFGLSGVGEPOYSRVRVFTTKAVAVITLDDTYDAY 306
Db 254 CRWKKDLDFQRLPYARDRVVGVFWISGVYFEPQYSLGRKMLTKVIAMASIVDDTYDSY 313
Qy 307 GTYEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTMEEFLLAKEGRDILFNGKE 366
Db 314 ATYEELIPYTNARWDIKCIDPEYMKPSYKALLDVYKEMEQSLVAEHGROYRVEYAKN 373
Qy 367 FVKEFVNLVMEAKWANEHGIPTTEBHDVPVITGGANLLTTTCYLGMSDIFTKESVEWA 426
Db 374 AMIRLAQSVLVEARWTLQNYKSFEEFKANALPTCGYAMLATISFVGMGDIPTPEFKWA 433
Qy 427 VSAPPLFRYSIGILGRRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEVAQTLIYKEVED 486
Db 434 ANDPKIIQASTIICRFMDVVAEHKFKHRRDDCSAIECYMEYGVTAQAEYDVFNKHVES 493
Qy 487 VKKDINREYLTNNIPRPLMAVILYQCLEVOYAGKDNFTMGDEYKHILKSLIYVPM 546
Db 494 AWKDNQEFKQKTEPTEVNLRLNRLARVMDVLYREGDGYTYVYVGAAGKGTISLLIEPIA 553
Qy 547 I 547
Db 554 L 554

RESULT 3

DCS2 GOSAR
ID DCS2 GOSAR STANDARD; PRT; 554 AA.
AC Q39760;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE (+)-delta-cadinene synthase isozyme XC14 (EC 4.2.3.13) (D-cadinene synthase).

QY 3 ALTEKPIRPTANFPSPISWGDOFLY-QKQVEGCVQIIVNDLKKEVQRLKZALDIPMKH 61
 Db 15 AISEN--RPKADFPFGIWDGMEIICPDIDAATELOQEELKAQVRKIMPEVD---DS 69
 QY 62 ANLLKLIDEIQRGLGIPYFEREIDHALQCIYETYGNNMGD-----RSSLFRLMRKOG 115
 Db 70 NOKLPFIADVQLGVSYHFEKIELENIYR---DTNNDADTDLYTALRFLREHG 126
 QY 116 YVTCVDFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLSM 175
 Db 127 FDISCAFNKFKDEAGNFKASLTSVQGLLELYEASVMVHGEDIIDEAISPTTAQLTLA 186
 QY 176 TKDAFSTPALFTEIQRALKQPLMKELPRIEAAQYIPFYQQODSHNKITLLKLAKLEFNLL 235
 Db 187 LP---TLHHPUSEQVGHALKQIRGLPRVEARNFISIQDLESHNKSLLQAKIDFNLL 243
 QY 236 QSLHKEELSHYCKWMAKFDIKKAPCLDRIVCEYFWGLSGVPEPOYSRVRVFTKAVAV 295
 Db 244 QLLHKEELSEICRWKWDIDFTKLPARDRVVEGVFWINGVFEPOYSGRKMLTKVIAM 303
 QY 296 ITLIDDTYDAYTYBELKIFTEAVERWSITCLDTLPEYMKPIYKUFMDTYTEMEEF 355
 Db 304 ASIVDDTYDSYATYDELIPYTAIERWDIKMQLPNYMKISYKALLNVYEEMEQLLAQ 363
 QY 356 GRTDLFNGCKEYFVKNLVEAKWANECHPTTEHDPVVIITGGANLLTTTCVLGMS 415
 Db 364 GRQVREYAKAMITRLQAYLEAKWTHQNYKPTFEFRDNALPTSGYAMLAITAFVGMG 423
 QY 416 DIFTKESVWAVSADPLFRYSIGILGRRLNDLMTKAEQKHSSSLESYMKYNNVEEY 475
 Db 424 EVITPETFKAASDPKIKASTIICRFMDDIAEHKFNHRRDDCSAIECYMKQYGTAGE 483
 QY 476 AQTLYKEVDVWKDINREYLTKNIPRLPLMAVLYLOFLEQVYAGKDNFTRMGDEYKH 535
 Db 484 AYNSFNKHIESWKDVNEEFKPTTEMPVLCSRLNARVMDVLYREBGDGYTHVGAAG 543
 QY 536 LKSLVYPMISI 547
 Db 544 GITSLLDIPQI 555

RESULT 5

DCSL_GOSHI STANDARD; PRT; 554 AA.
 AC P93665;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE (+)-delta-cadinene synthase (EC 4.2.3.13) (D-cadinene synthase).
 GN CDNL.
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID=3635;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Westburn M;
 RA Davis E.M., Chen Y.-S., Essenberg M., Pierce M.L.;
 RT "cDNA sequence of a (+)-delta-cadinene synthase gene induced in
 RT Gossypium hirsutum L. by bacterial infection."
 RL (in) Plant Gene Register PGR98-040.
 RN [2]
 RN PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RP MEDLINE=96351891; PubMed=8728715;
 RX Davis E.M., Tsuji J., Davis G.D., Pierce M.L., Essenberg M.;
 RA "Purification of (+)-delta-cadinene synthase, a sesquiterpene cyclase
 RT from bacteria-inoculated cotton foliar tissue."
 RL Phytochemistry 41:1047-1055(1996).
 CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
 CC diposphate (FPP) to (+)-delta-cadinene.
 CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diposphate = (+)-

CC delta-cadinene + diphosphate.
 CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
 CC first (committed) step.
 CC -!- INDUCTION: By bacterial infection.
 CC -!- SIMILARITY: Belongs to the terpene synthase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U88318; AAC12784.1; --
 DR HSP; Q40577; SEAU.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR001906; Terp_synth-like.
 DR InterPro; IPR005630; Terpene_synth_C.
 DR InterPro; IPR008949; Terpenoid_synth.
 DR Pfam; PF01397; Terpene_synth_1.
 DR Pfam; PF03936; Terpene_synth_C; 1.
 KW Lyase.
 FT ACT_SITE 451 451 BY SIMILARITY.
 FT ACT_SITE 527 527 BY SIMILARITY.
 FT ACT_SITE 531 531 BY SIMILARITY.
 SQ SEQUENCE 554 AA; 64019 MW; 8BCC78AD8CA5B816 CRC64;
 Query Match 38.7%; Score 1126; DB 1; Length 554;
 Best Local Similarity 41.4%; Pred. No. 3,3e-69;
 Matches 224; Conservative 111; Mismatches 196; Indels 10; Gaps 5;
 QY 10 IRPIANPPISWGQFL-IYQKVEGCVQIIVNDLKKEVQRLKZALDIPMKH-LKL 67
 Db 21 MRPKADFPQISWGDFLNCPPDKNIDAEQKHHQQLKEEVKQIV---APMANSTKLAF 76
 QY 68 IDEIQSLGIPYHFEREIDHALQCIYETYGNNMGD-DRSSLFRLMRKOGYVTCVDFNNY 126
 Db 77 IDSVOGLGVSYHFTKELEDELENIYHNNDALDLYTTSRFLRLREHGFHSCDVFNKF 136
 QY 127 KDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRRLSINTKDAFSTNPA 186
 Db 137 KDEQGNFKSVTSDVRGLLELYQASVLRVHGEDIIDEAISFTSNHLSLAVA---SLDHPL 193
 QY 187 FTEIQALKQPLMKRLPRIEAAQYIPFYQQODSHNKITLLKLAKLEFNLLQSLHKEELSHV 246
 Db 194 SEVSHALKQSIRRGLPVREARHVLVYQDIESHNKVLLEFAKIDFNVMVQLLHRELSEI 253
 QY 247 CKWKAFADIKKAPCLDRIVCEYFWGLSGVPEPOYSRVRVFTKAVAVITLIDDTYDAY 306
 Db 254 SRWKDLDFQKLPYARDRVVEGVFWISGVYFEPQYSLGRKMLTKVIAMASIVDDTYDSY 313
 QY 307 GTVEELKIFTEAVERWSITCLDTLPEYMKPIYKUFMDTYTEMEEFLEKEGRTDLFNCCKE 366
 Db 314 ATVEELIPYTAIERWDIKIDELPEYMKPSYKALLDVEEMEQLVAEHGQYRVEYAKN 373
 QY 367 FVKEFVRNLMVEAKWANECHPTTEHDPVVIITGGANLLTTTCVLGMSDITFKESVMA 426
 Db 374 AMIRLAQSVLVEARWTLQNYKPSFEPEFKANALPTCGYAMLAITSFVGMGDIIVTPEFKWA 433
 QY 427 VSAPPLFRYSIGILGRRLNDLMTKAEQKHSSSLESYMKYNNVEEYAOYTLIYKEVD 486
 Db 434 ANDPKIIQASTIICRFMDDVTTEHKFKHRRDDCSAIECYMEBYGVTAQEAIVDFNKHVES 493
 QY 487 VWKIDINREYLTKNIPRLPLMAVLYLOFLEQVYAGKDNFTRMGDEYKHLIKSLVYPMIS 546
 Db 494 AKWDVNGQFLKPTTEPTEVLNRSINLARVMDVLYREGDGYTVVGKAAKGITSLTLLIEPIA 553
 QY 547 I 547
 Db 554 L 554

QY	6	EEKPIRPIANFP	34
Query Match	34		
Best Local Similarity	39		
Matches	221	Conservative	
FT	HELIX	36	5
FT	TURN	58	58
FT	TURN	60	61
FT	HELIX	64	76
FT	TURN	77	78
FT	HELIX	80	83
FT	HELIX	84	97
FT	HELIX	104	116
FT	TURN	117	118
FT	HELIX	123	129
FT	STRAND	130	130
FT	TURN	132	133
FT	STRAND	136	136
FT	HELIX	138	142
FT	HELIX	144	154
FT	TURN	155	156
FT	TURN	159	160
FT	HELIX	162	164
FT	TURN	165	166
FT	HELIX	167	178
FT	HELIX	179	181
FT	TURN	184	185
FT	HELIX	186	195
FT	TURN	199	200
FT	HELIX	203	213
FT	TURN	214	214
FT	HELIX	215	217
FT	TURN	219	220
FT	HELIX	223	254
FT	TURN	255	255
FT	HELIX	256	259
FT	TURN	261	262
FT	HELIX	267	277
FT	HELIX	281	283
FT	HELIX	284	306
FT	TURN	307	307
FT	HELIX	310	322
FT	TURN	323	323
FT	HELIX	326	330
FT	HELIX	333	354
FT	TURN	355	357
FT	HELIX	359	361
FT	HELIX	362	385
FT	TURN	386	386
FT	HELIX	391	398
FT	TURN	399	399
FT	HELIX	400	402
FT	HELIX	404	413
FT	TURN	414	414
FT	TURN	416	417
FT	HELIX	420	427
FT	TURN	428	428
FT	HELIX	431	454
FT	TURN	455	456
FT	TURN	458	459
FT	HELIX	461	469
FT	TURN	470	470
FT	HELIX	473	494
FT	TURN	495	495
FT	HELIX	503	505
FT	TURN	506	506
FT	HELIX	507	519
FT	TURN	520	520
FT	HELIX	535	542
FT	TURN	543	543
SQ	SEQUENCE	548 AA;	629

Query Match 34.9%; Score 1015.5; DB 1; Length 548;
Best Local Similarity 39.8%; Pred. No. 1.1e-61;
Matches 221; Conservative 102; Mismatches 203; Indels 29; Gaps 9;

Db 10 EBEIVRPVADPSLWGDQFLSFIDNQVAKVAKETALKEQTRNML---LATGMKLD 66

Qy 64 LLKLIDSIQRLGTPYHFEREDHALQCIYIYEGDNWGD-----RSSLWRLMRKQYV 118

Db 67 TUNLIDTIERLGSYHFEKIDBILDQIY-----NQNSNCNDLCTSLAQFRLRHQGFNI 121

Qy 119 TCDVFNKYDKNGAFKQSLANDVEGLLELYEATSMRVPGHIILEDALGFTSRSLSIWKD 178

Db 122 SPEIFKFDENGKFKESLASDVLGLNLVEASHVRTHADILEDALAFSTHLE----- 176

Qy 179 APTNPAFT-----ETQALQKPLMKRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233

Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFISSTYDKESQKNVLLRFKLDEN 234

Qy 234 LLOSLKESLHVCKWKAFAIDIKKNAFCPLDRIVECYFWGLSGSYEPOYSRPFVKAV 293

Db 235 LQMLKHEQEAQVSRWKOLDFTTLPYADRVRVCEYFWALGVYFEPYSOARVMLVKTI 294

Qy 294 AVITLIDDTYDAYGYVEELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEBFLA 353

Db 295 SMISIVDDIFDAGTVKLEAYDAICQWIDNEIDRLPDYMKISYKAILDLYKDYKELS 354

Qy 354 KEGRTDLNCGKEFVKFVFNLMVEAKWANEHPIPTTEHDHPVVIITGGANLLTTTCYLG 413

Db 355 SAGRSHIVCHAIERMKEVVRNVESTWFTFEGYTPPVSEVLSNALATTYYVLLATTSYL 414

Qy 414 MSDIFTKESVWAVSAPPLFRYSGLGRRLNDLMTKAEQERKSSSLESYMEYVNE 473

Db 415 MKSA-TEQDFEWLSKPKILEASVILICRVDDTATYEVKSRGQIATGIBCCMDYGLST 473

Qy 474 EYAQTLLYKEVEDVWKDINREYLTNNIPRLMAVILYQFLVEQVA-GKDNFTRMGDE 532

Db 474 KEAMAFQNMATKWDINGELRLPTVPVSTFTPLNLARIVETVIHNLGTYTHPEKV 533

Qy 533 YKHLIKSLLYPMSI 547

Db 534 LKPHIINLLVDSIKI 548

RESULT 7

ID CASS_RICCO STANDARD; PRT; 601 AA.

AC P5287;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Casbene synthase, chloroplast precursor (EC 4.2.3.8).

OS Ricinus communis (Castor bean).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;

OC eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;

OC Ricinus.

OX NCBI_TaxID=3988;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94359958; PubMed=8078910;

RA Mau C.J., West C.A.;

RT "Cloning of casbene synthase cDNA: evidence for conserved structural

RT features among terpenoid cyclases in plants.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:8497-8501(1994).

RN [2]

RP CHARACTERIZATION.

RX MEDLINE=97115639; PubMed=8954576;

RA Hill A.M., Cane D.E., Mau C.J., West C.A.;

RT "High level expression of Ricinus communis casbene synthase in

RT Escherichia coli and characterization of the recombinant enzyme.";

RL Arch. Biochem. Biophys. 336:283-289(1996).

CC -!- FUNCTION: Catalyzes the cyclization of geranylgeranyl diphosphate

CC to casbene, a diterpene phytoalexin with antibacterial and

CC antifungal activity.

CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = casbene +

CC diphosphate.

CC -!- SUBCELLULAR LOCATION: Chloroplast.

CC -!- INDUCTION: By oligogalacturonide fragments released by fungal

CC infection. Detected after 5 h of incubation with the pectic

CC fragments and reaches a maximum after 10-12 h.

CC -!- MISCELLANEOUS: The Km of this enzyme is 1.9 micromol.

CC -!- SIMILARITY: Belongs to the terpene synthase family.

CC -!- CAUTION: It is uncertain whether Met-1 or Met-8 is the initiator.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-eib.ch/announce/>

CC or send an email to license@isb-eib.ch).

CC -----

CC EMBL; L32134; -; NOT ANNOTATED CDS.

DR InterPro; IPR008930; Terp_cyc_toroid.

DR InterPro; IPR001906; Terp_synth-like.

DR InterPro; IPR005630; Terpene synth C.

DR InterPro; IPR008949; Terpenoid synth.

DR Pfam; PF01397; Terpene synth; I.

DR Pfam; PF03936; Terpene synth C; 1.

KW Plant defense; Lyase; Chloroplast; Transit peptide.

FT TRANSIT 1 56

FT CHAIN 57 601

FT ACT_SITE 500 500 BY SIMILARITY.

FT ACT_SITE 575 575 BY SIMILARITY.

FT ACT_SITE 579 579 BY SIMILARITY.

SQ SEQUENCE 601 AA; 68965 MW; F7B32D286747957 CRC64;

Query Match 33.3%; Score 967.5; DB 1; Length 601;

Best Local Similarity 35.9%; Pred. No. 2.e-58;

Matches 199; Conservative 120; Mismatches 220; Indels 15; Gaps 6;

Qy 2 MALTEKPIRPIANPPPSINGDQF--LIYQKQVQGVQIIVNDLKKEVQLLKEALDIPM 59

Db 55 LSSITTHQEVRLPAYPPPTWGNRFASLTNPSEFESYDERVIVLKKVKDIL---ISSTS 111

Qy 60 KHANLLKLIDSIQRLGIPYHFEREDHALQCIYIYEGDNWGDSSLW-----FRLMRKQ 114

Db 112 DSVETVILIDILLCRLGVSYHFEKIDBILDQIY-----NQNSNCNDLCTSLAQFRLRHQGFNI 171

Qy 115 GYVTCDFVNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGHIILEDALGFTSRSLSI 174

Db 172 GFKMSDVSFKSDGKSGKSLRGDAKGLSLFASHLSVHGEDILEAFATKYQLS 231

Qy 175 MTKDAFSNPALFTFIQKALKQPLMKRLPRIEAAQYIPFYQQOQ-DSHNKTLLKLAKLEFN 233

Db 232 SAVELF---PNLKRHTNALQEPFHSGVPRLEARKFIDLYEADIECRNETLLEFAKLDYN 288

Qy 234 LLOSLKESLHVCKWKAFAIDIKKNAFCPLDRIVECYFWGLSGSYEPOYSRPFVKAV 293

Db 289 RVQLLHQEQLCOFSKWKNDNLASDIPYARDMAEIFFWAVAMYFFDFVAHTRMIAKV 348

Qy 294 AVITLIDDTYDAYGYVEELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEBFLA 353

Db 349 LLISLIDDTIDAYATWEETHILAEAVARDMSCKLEKPDYMKVIYKLLNTTFSEFEKELT 408

Qy 354 KEGRTDLNCGKEFVKFVFNLMVEAKWANEHPIPTTEHDHPVVIITGGANLLTTTCYLG 413

Db 409 AEGKSYVKYGREAFQELVRGYLEAVWRDEGKIPSFDDLYNGSMTTGLPLVSTASFMG 468

Qy 414 MSDIFTKESVWAVSAPPLFRYSGLGRRLNDLMTKAEQERKSSSLESYMEYVNE 473

Db 469 VQETGLNFEQWLETNPKLSYASGAFIRLVNLDLTSHVTEQQRGHVASCIDCYMNHGVSK 528

Qy 474 EYAQTLLYKEVEDVWKDINREYLTNNIPRLMAVILYQFLVEQVA-GKDNFTRMGDE 533

Db 529 DEAVKILQKMATDCWKEINEECRQSQSVSGHLMRIVNLARLTDVSYKYGDTD-SQQL 587

Qy 534 KHLIKSLLYPMSI 547

Db 588 KQFVKGLFVDPISI 601

```

170 VTDLNSTALGLRLTLRLHGYPVSSDVFKAFKQNGQFSCSENIQTDEIRGVNLFRASLI 229
154 RVPGEIILEDALGFTSRSLRSLNMTKDAFSTNP--ALFTIQRALKQPLWKLPRIERAAQVI 211
230 AFPEKINDAEAFSTKYL-----KEALQKIPVSLRSGIDVLEYGWHTYLRLEARNYI 285
212 PFYQQQDSHN-----KTLKLAKLEFNLLQSLHKEELSHVCKWKKAFDIKKNAPCLRD 264
286 QVF-GQDTENTKSVKSKULELAKLEFNIFQSLQKRELSIVRWKSGFPEMTFC-RH 343
265 RIVCEYFWGLSGVEPOYSRARVFTTAKAVITLIDDDYDAYGTVEELKIFTEAVERWSI 324
344 RHVEYVTLASCIAPFQHSGLFGLFAGTKCHLITVLDDMYDTFTGTVDELELFATMKRWD 403
325 TCLDTLPEYMKPKYKLFMDIVTEM-BEFLAKEGRTDLFNCGKGFVKFVRNLMVEAKWAN 383
404 SSIDCLPEYMKGVIAVYDTVNEMAREAEAAQGR-DTLTYAREAEVADISYMQEARWIA 462
384 EGHPTTEEHDPVVIIITGGANLLTTTCYLGMSDIFTKESVEMAVSAPPLFRYSGILGREL 443
463 TGYLPSFDEYENGKVSCHGRISALQILTM-DIPFPDHLKEVDFF-----SKL 511
444 NDL-----MTHKAEQERKSSSSLESYKEY-VNNEEYAOQLIYKEVEDVWKDI 491
512 NDLACAILRLRGDTRCYKADRGEEASS-SCYMKDNPVSEEDALDHINAMISDVIKGL 571
492 NREYIATK-NIPRLMAVYLCQFLEVQAGKDNFTRMGDEYKHLIKSLIV 542
572 NWEILLKPDINVPISAKKHAFDIARAFHYGYKRDGYSVANVETKSLVTRTL 623

RESULT 9
TSD2_ABIGR
ID TSD2_ABIGR STANDARD; PRT; 627 AA.
AC 024474;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myrcene synthase, chloroplast precursor (EC 4.2.3.15).
GN AG2.2.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
NCBI_TaxID=46611;
RX [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE=97413772; PubMed=9268308;
RA Bohmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT characterization, and functional expression of myrcene synthase, (-)-
RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase."
RL J. Biol. Chem. 272:21784-21792(1997).
CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC response to insect attack or other injury. Involved in monoterpene
CC (C10) olefins biosynthesis.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate = pinene + diphosphate.
CC -!- COFACTOR: Manganese and potassium.
CC -!- PATHWAY: Oleoresinosis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- INDUCTION: By wounding.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U87908; AAB71084.1; -
CC HSSP; Q40577; SEAS.
CC InterPro; IPR008930; Terp_cyc_toroid.

```



```

QY 389 TTEHDPVVIITGGANLLTTTCYLGMSDIFTKESVEMVAVSAPPLFRYSGLRRLNDLMT 448
Db 702 TPEEYLKTVIAISVGLGPCLOPILLMGLVLDVVEKVVHPSNMFELVLSWLTNDTKT 761
QY 449 HKAEERKHSLSLSYKEY-NVNEEYAQTLLIYKEVEDVWKDINREYLTQX-IPRPLL 506
Db 762 YQAEKARQQAGGIACYNKDNFPCATEADAIGHICRVVDRAKLEASPEYFKPSNDIPMGCK 821
QY 507 MAVIYLCQFLEVOYAGKNFTRMGDYKHLIKLSLLVYPMXI 547
Db 822 SFIFNLRLCVQIFYFIDGYGIANEIKDYIRKVYIDPIQV 862

RESULT 11
TASY TAXBA
ID TASY TAXBA STANDARD; PRT; 862 AA.
AC Q93V33
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN TDC1 OR TASY.
OS Taxus baccata (English yew).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX NCBI_TaxID=25629;
RN [1]
RP SEQUENCE FROM N.A.
RA Goerhardt B.;
RL Thesis (2001), Technische Universitaet Berlin, Germany.
CC -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC parent olefin with a taxane skeleton.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC diphosphate.
CC -!- PATHWAY: Taxol biosynthesis; first step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ320538; CAC42773.1; -
CC InterPro; IPR008930; Terp_cyc_toroid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene synth C.
CC InterPro; IPR008949; Terpenoid synth.
CC Pfam; PF01397; Terpene_synth_1.
CC Pfam; PF03936; Terpene_synth_C; 1.
CC KX Taxol biosynthesis; Lyase.
CC ACT SITE 758 758 BY SIMILARITY.
CC ACT SITE 835 835 BY SIMILARITY.
CC ACT SITE 839 839 BY SIMILARITY.
CC SEQUENCE 862 AA; 98048 MW; B37835AFFCIDEFE CRC64;

Query Match 20.7%; Score 601.5; DB 1; Length 862;
Best Local Similarity 29.9%; Pred. No. 2.8e-33;
Matches 163; Conservative 89; Mismatches 232; Indels 61; Gaps 13;

QY 50 LKREALD-----IPMKHA-----NLKLDIEIQRGLPIYHFEIDHALOCIVETY---GD 97
Db 332 LLNNLLDFGCGVPCWYSIDLRLSLVDNIEHLGIRHFKQBIKVALDYVYVRHWSERGI 391
QY 98 NWNGD-----RSLSWFLMRKQGYVTCDFVNNYKDKNGAFKQSLAN---DVEGLLEL 147
Db 392 GWGSDSLVDPDNLNTALGLRLTLTHGYDVSSDVLNNPKDENGFFSAGOTHVELRSVNL 451
QY 148 YEATSMRVFGEIILEDALGFTSRSLISINTKDAFTSNPALFTIQRALKQPLWRLPRIA 207

```

```

Db 452 FRASDLAPPDEGAMDARKPAEPYLRDALATKISTNTKLYKEIYVVVEPWHMSIPRLEA 511
QY 208 AQYI-----PPYQQODSHNKTLKAKLEFNLLQSLHKEELSHVCKWKAFDI 255
Db 512 RSYIDSVDVVWOKTLYRMPSLSNCKLELAKLDFNIVQSLHOELKLLTRWKESGM 571
QY 256 KQNAPCLRDRIVECYFWMGLSGYPQYSRARVFFTKAVAVITLIDDTYDAYGYEELKIF 315
Db 572 -ADINFTRHRAVEYVF--SSATFPEYSATIAFTKIGCLOVLFDMDADIATLDELKSF 628
QY 316 TEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLEAK-EGRTDL-----FNCG 364
Db 629 TEGVKRWDTSLLHEIPECMQTCFKVWFKLMEVNDVVKVQGRDMLAHIRKPEWELYNFC- 687
QY 365 KEFVKFVRNLWKAWEKANEHGIPTTEBHDPIVITGGANLLTTTCYLGMSDIFTKESVE 424
Db 688 -----YVQERWLEAGYIPTFEYLYKTYALSGLPCTLOPILLMGLVLDVVE 737
QY 425 WVSAPPLFRYSGLRRLNDLMTKAEQERKHSLSLSYKEY-NVNEEYAQTLLIYKE 483
Db 738 KVHYPNMFELVLSWLTNDTKYQAEKARQQAGGIACYNKDNFPCATEADAIGHICRV 797
QY 484 VEDVWKDINREYLTQX-IPRPLLMAVYLCQFLEVOYAGKNFTRMGDYKHLIKLSLV 542
Db 798 VDRALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIFYFIDGYGIANEIKDYIRKVYI 857
QY 543 YPMXI 547
Db 858 DPIQV 862

RESULT 12
TASY TAXCH
ID TASY TAXCH STANDARD; PRT; 862 AA.
AC Q9PT37;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN TDC1.
OS Taxus chinensis (Chinese yew).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX NCBI_TaxID=29808;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Callus;
RC Wang W., Shi Q., Ouyang T., Zhu P., Cheng K.;
RT "Cloning, expression, and characterization of taxadiene synthase, a
RT diterpene cyclase from Taxus chinensis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC parent olefin with a taxane skeleton.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC diphosphate.
CC -!- PATHWAY: Taxol biosynthesis; first step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY007207; ARG02357.1; -
CC HSPP; Q40577; SEAU.
CC InterPro; IPR008930; Terp_cyc_toroid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene synth C.
CC InterPro; IPR008949; Terpenoid synth.

```

DR Pfam; PF01397; Terpene synth; 1.
 DR Pfam; PF03936; Terpene synth_C; 1.
 KW Taxol biosynthesis; Lyase.
 FT SITE 758 758 BY SIMILARITY.
 FT ACT SITE 835 835 BY SIMILARITY.
 FT ACT_SITE 839 839 BY SIMILARITY.
 SQ SEQUENCE 862 AA; 98069 MW; 3A597BAF72B2BF679 CRC64;
 Query Match
 Best Local Similarity 30.6%; Score 598; DB 1; Length 862;
 Matches 159; Conservative 85; Mismatches 225; Indels 52; Gaps 11;
 QY 65 LKLDLQRIQIGYPHFREIDHALQCIYETV---GDNWNGD-----RSLPFLMRKQ 114
 DB LSLVDNTEHLGIGRHFQKQKVALDYVYRWSESGIGWGRDLSLPDLMTTALGLRLTRTH 415
 QY 115 GYVVTCDVFNKYDKNGAFKQSLAN---DVEGLLELYEATSMRVPGLEILLEDALGFTSR 171
 DB GYVSSDLNMFNDENGRFFSSAGQTHVELRSVVILFRASDLAPPDEGAMDDARKFAEPY 475
 QY 172 LSIWTKDAFSTNPALFTEIORALKQPKWLKPRIEAAQYI-----PFYQQQDS 219
 DB LRDALATKISTNYKLFKEIEVVEYPMHGIIPRSEARSYIDSDYDDYVWERKTLRNP 535
 QY 220 HNTKLLKAKLEPNLQSLHKEELSHVCKWKWAFDIKKNAPCLDRIVECYFWGLGSGYE 279
 DB SNGKLELAKLDFNIVQSLHQBELKLLTRWKKESGM-ADINFTRHRAEVVF--SSATFE 592
 QY 280 PQYSRARVFTKAVAVITLDDYDYAGTYEELKIPTFAVERWSITCLDILPEYKPIYK 339
 DB PEYSATRIAEFTKIGCLQVLFDDWADIFATILDELKSFTEGVKRWDTSLHRIPECMQTCFK 652
 QY 340 LFMVDTYMEEFELAK-EGRTDL-----FNCQGEFKEFVRNLMVAKWANEHGP 388
 DB VWFKLIEVNDVVVQGRDMLAHIRKFWELYNFC-----YVQREWLDAIYIP 701
 QY 389 TTEHDPVWITGGANLLTTTCVGLMSDIFTKESVEWAVSAPPLFRYSYGILGRDLNMT 448
 DB TFEYLKTYAISVLGPCTIQLPILLMCELKVDVVEKVVHVPNSNMFLVSLSWRLTNDTKT 761
 QY 449 HKAEQKHSKSSLESYMKK-YNNVEYEAQTLIYKEVEDWVKDINRILYITKN-IPRLL 506
 DB YQAEKARGQOASGIACMKDNLGATEBDAIKHCRVVDRAKKEASFEYKPSNDIPGCK 821
 QY 507 MAVYLICQFLEVQYAGKDNFTMGDEYKHLIKSLLVYPM 547
 DB SFINRLVCQIFKFIKFDIGGIANEELKXIRKVIYDPIQV 862
 RESULT 13
 ID TSD3_ABIG STANDARD; PRT; 637 AA.
 AC Q22340; 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE (-)-(4S)-limonene synthase, chloroplast precursor (EC 4.2.3.16).
 GN AG10.
 OS Abies grandis (Grand fir).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
 ON NCBI_TaxID=46611;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RA MEDLINE=97413772; PubMed=9268308;
 RX Bohlmann J., Steele C.L., Croteau R.;
 RT "Monoterpene synthases from grand fir (Abies grandis): cDNA isolation,
 RT characterization, and functional expression of myrcene synthase, (-)-
 RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
 RL J. Biol. Chem. 272:21784-21792 (1997).
 CC -1- FUNCTION: Involved in defensive oleoresin formation in conifers in
 CC response to insect attack or other injury. Involved in monoterpene
 CC (C10) olefins biosynthesis.

[illegible]

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE016967; AAP56564.1; -
DR EMBL; L38402; AAB40952.1; -
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_A.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF06623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SM00663; RPOLA_N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription;
KW Complete proteome.
SQ SEQUENCE 1286 AA; 145058 MW; FB8C242C2D4AC693 CRC64;

Query Match 4.3%; Score 124; DB 1; Length 1286;
Best Local Similarity 20.7%; Pred. No. 1.4;
Matches 133; Conservative 89; Mismatches 194; Indels 226; Gaps 37;

QY 2 MALTTE-KEIRPIANPPSIWGQFLIYQKVEGQVQVNDLKEVRQLLKEALDIPMK 60
Db 445 MFNEDGGQIKPIA---PHIKIAQMIQAQ--SERIWDIVDKVIKERPVULNRA---PTL 496
QY 61 HANLLKLIDEIQRIGIPYHFERID-----HALQCIYETYGDNMNGDRSL----- 106
Db 497 H-----RLGIQAEPKIVDGAIRLHPL--VTTAFNADFDGDQMAVHPISKEA 543
QY 107 -----WFRMRKQGYV---TCD-VFNYY-----KDKNGAFQSLANDVEGLLEL 147
Db 544 VAERAIMLASWHILGPDKGPVATPDQMVGLGNVYLTTEKRNEKGEGLIFSDPDQVILA 603
QY 148 YEATSRVPEIILEDALGFTSRSLMTKDAFTNPALFTEIQRALKQPLWKELPRIEA 207
Db 604 YEAKQVSIHALIGLS-----TKCLTKKPKAQGIVITVTKAIMNSI---MP--EE 649
QY 208 AQYIPFYQQQDSHNTKLL-----KLAKLEFNLLQSLKKEELSHVCK-WWKA 252
Db 650 MAYL-----NDGNLLELDESIVFAGEDPKQKLAKRP--LYKPPGKKTLSKIEILYKN 702
QY 253 FDIKNAPCLRDRIVECYFWGLSGYE-----PQYSRVRVFTKAVAVITLID 300
Db 703 FPLQK-VPOVLDKIKE-----FGFKYSTLSSTTISVFDIPRYDNKQEIYITKANEMIAKLK 756
QY 301 DTYDAVGYEELKIPTFAVERWS-----ITCLDTLPEYMKPIYKLFMDT----- 344
Db 757 HMYQK-GLTDDERYTKVRLNADVNDRDIKEITRPEYKENSIVVIADSGARGNIS 815
QY 345 -YTE---MEEFLLAKEGRTDLFCGKEFVKFEFVRNLMVEAKWANEHGIPTTEEDPVIIT 400
Db 816 NTFQFGMRGLMSKS-----YNYDQIKSQVIRDTI-----EVP1--KHS---FIE 856
QY 401 GGANLITTCYL-----GMSDIFTKESVEWAVSAPPLFRYSGILGRINDLMTKAE 452
Db 857 G-----LTINEYFNSSYGARKGMTDIAMKTS-----KSGVMTRKLVD-----AA 895
QY 453 QERKHSSSL-----ESYMKVYVNEEY-----AQT 478
Db 896 QEVINDSCNTNKGIVVSTITNSLDGGVVTLSERIVTRYITDIPYDEKTKELLVDADT 955
QY 479 LIYKEVEDVMKIDINREYLTNIPRLLMAVIY-----LCQ 514
Db 956 LITSLE-----AEKIAKANVTKALIRSPIYQSTKGLCQ 989

Search completed: June 3, 2004, 16:39:28
Job time : 24 secs